약학박사학위논문

가자나무 열매에서 분리한 알파 글루코시데이즈 억제 성분

α-Glucosidase Inhibitory Constituents from *Terminalia chebula* Fruits

2017 년 2 월

서울대학교 대학원

약학과 생약학전공

이 동 영

Abstract

Type 2 diabetes mellitus (T2DM), known as non-insulin-independent diabetes, is a chronic metabolic disorder which is due to insulin-resistance in the tissues. As a result, insulin resistance leads to elevated blood glucose levels, which can damage many of the organs. One of the therapeutic approaches for treating T2DM is to alleviate postprandial hyperglycemia. This is achieved by suppressing the glucose absorption from the gut by inhibiting intestinal carbohydrate digesting enzymes such as α-glucosidase. In this study, to isolate new α-glucosidase inhibitory compounds efficiently, an HPLC-based activity profiling with dereplication was carried out. HPLC-based activity profiling of an extract from *Terminalia chebula* fruits enabled the isolation and identification of thirteen compounds including four new ones (**3** and **11**-**13**). Among the thirteen compounds, compounds **4**-**13** obtained from the active microfractions exhibited potent inhibitory activities against Baker's yeast α-glucosidase. These results confirm that HPLC-based activity profiling is an effective method for isolating new bioactive compounds.

Moreover, to investigate the constituents of *T. chebula* fruits, further isolation was carried out. As a result, fifty-three compounds including thirty-six hydrolysable tannins and seventeen polyhydroxytriterpenes were further isolated from a methanolic extract of *T. chebula* fruits. A total of sixty-six compounds, including nine new hydrolysable tannins (**3**, **11**-**13**, **18**, **20**, **21**, **23**, and **24**) and three new polyhydroxytriterpene derivatives (**51**, **52**, and **66**), were finally isolated and identified.

The inhibitory activities of all the isolated compounds against Baker's yeast α-

glucosidase were tested, and as a result, twenty-four compounds (**4**-**13**, **16**-**18**, **22**, **25**, **43**-**49**, **60**, and **61**) exhibited inhibitory activities. Among these compounds, compounds **8**, **11**, and **12** showed significant inhibitory activity (IC_{50} 8.3, 6.4, and 2.9 μM, respectively). Furthermore, all the compounds were tested for their inhibitory activity against rat intestinal *α*-glucosidase, porcine pancreatic *α*-amylase, and PTP-1B. In the case of rat intestinal *α*-glucosidase and porcine pancreatic *α*-amylase, compound **8** showed the potent inhibitory activity (IC_{50} 17.3 and 13.4 μM, respectively) which was comparable to the positive control, acarbose. In the case of PTP-1B, compounds **8** and **60** exhibitied the most significant inhibitory activities (IC_{50} 2.0 and 10.3 μM, respectively) among the hydrolysable tannins and polyhydroxytriterpene derivatives, respectively.

Key words: *Terminalia chebula*, Combretaceae, hydrolysable tannin, polyhydroxytriterpene, HPLC-based activity profiling, *α*-glucosidase inhibitory activity

Student Number: 2009-21699

Contents

List of schemes

List of tables

List of figures

List of abbreviations

CC: column chromatography

$CHCl_3$: chloroform

cheb: chebuloyl

Cin: cinnamoyl

COSY: correlation spectroscopy

d: doublet

DMSO: dimethyl sulfoxide

DTT: dithiothreitol

ESIMS: electrospray ionization mass spectroscopy

Gal: galloyl

Glc: glucose

HHDP: hexahydroxydiphenoyl

HMBC: heteronuclear multiple bond correlation

HPLC: high performance liquid chromatography

Hz: hertz

IC_{50}: the half maximal inhibitory concentration

IR: infrared radiation

m: multiplet

MeCN: acetonitrile

MeOH: methanol

n-BuOH: n-butanol

neocheb: neochebuloyl

NMR: nuclear magnetic resonance

ODS: octadecylsilane

PBS: phosphate-buffered saline

Rha: rhamnose

RP: reverse phase

s: singlet

t: triplet

t_R: retention time

UHPLC-QTOF-MS: ultra-high performance liquid chromatography quadrupole time-of-flight mass spectrometry

UV: ultraviolet

I. Introduction

Terminalia chebula Retz.

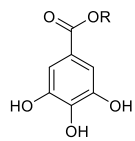
Terminalia chebula Retz., which belongs to the family of Combretaceae, is widely distributed in India and Southeast Asia and extensively cultivated in Taiwan (Cheng et al., 2003). *T. chebula* is a medium-sized deciduous tree; reaching a height of up to 25 m with a broad disk-shaped crown and wide spreading branches (Anwesa et al., 2013). *T. chebula* is an important medicinal herb in traditional Indian medicine and it is the most frequently used herb in Ayurveda (Rathinamoorthy et al., 2014). In addition, the fruit of *T. chebula* has been regarded as the "King of the medicine" by Tibetans and held in high regard by fork medicine practitioners (Klika et al., 2004). The fruit of *T. chebula* has been commonly used in ethnomedicine due to its laxative, diuretic, homeostatic, antitussive, and cardiotonic properties (Barthakur and Arnold, 1991). The extracts of *T. chebula* have been found to possess various bioactivities such as antioxidant, hepatoprotective, cytoprotective, immunomodulatory, antimicrobial, radiation-protective, hypolipidemic, and antidiabetic activities (Table 1). The various type of secondary metabolites such as the hydrolysable tannins including simple gallotannins, ellagitannins, chebulic ellagitannins, ellagic acid derivatives and glycosides, phenolic acids, flavonoids, triterpenes, and their glycosides (Figure 1-3), have been reported from the *T. chebula* (Fahmy et al., 2015). Especially, *T. chebula* is known to be rich in hydrolysable tannins.

Table 1. Pharmacological studies reported in *T. chebula*

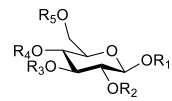
Therapeutic target	In vitro/vivo assay	Extracts/Active constituents	Reference
Anti-ulcerogenic	Aspirin, ethanol, cold restraint stress induced ulcer model	Methanol extract	(Raju, 2009)
Neuroprotective	PC_{12} cells	Methanol and aqueous extract	(Chang and Lin, 2012)
Antiviral	HIV-1 integrase	Gallic acid and gallotannins	(Ahn et al., 2002)
	Herpes simplex virus (HSV)	Aqueous extract	(Kurokawa et al., 1995)
	Human cytomegalovirus (CMV)	Aqueous extract	(Yukawa et al., 1996)
Anticonvulsant	Maximal electroshock induced seizure	Ethanol extract	(Debnath, 2010)
Antioxidant	HRP-luminol-H_2O_2 assay	Methanol extract	(Cheng et al., 2003)
Antidiabetic	Alloxan-induced diabetic rat	Hydroalcoholic extract	(Sabu and Kuttan, 2002)
Anticaries	Glucan-induced aggregation of *Streptococcus mutans*	Aqueous extract	(Jagtap and Karkera, 1999)
Hepatoprotective	Rifampicin, isoniazid induced hepatotoxicity	Hydroalcoholic extract	(Tasduq et al., 2006)
Hypolipidemic	Atherogenic diet induced hyperlipidemic rat	Hydroalcoholic extract	(Hasani-Ranjbar et al., 2010)
Immunomodulatory	Cytotoxic T lymphocyte mediated cytotoxicity	Gallic acid, chebulagic acid	(Hamada et al., 1997)
Radioprotective	Irradiation	Aqueous extract	(Jagetia et al., 2002)

Wound healing	Fibroblast, collagen synthesis	Hydroalcoholic extract	(Saha, 2011)

Gallic acid and gallotannins

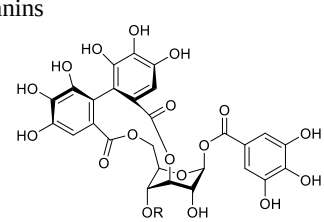


R=H: Gallic acid
R=Me: Methyl gallate

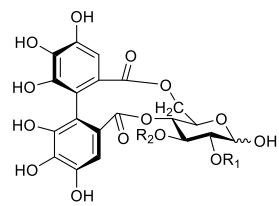


R_1=R_5=Gal, R_2=R_3=R_4=H: 1,6-Di-*O*-galloyl glucose
R_1=R_2=H, R_3=R_4=R_5=H: 3,4,6-Tri-*O*-galloyl glucose
R_1=R_3=R_4=R_5=Gal, R_2=H: 1,3,4,6-Tetra-*O*-galloyl glucose
R_1=R_2=R_3=R_4=R_5=Gal: 1,2,3,4,6-Penta-*O*-galloyl glucose

Ellagitannins

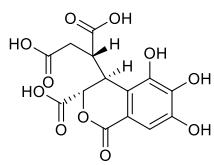


R= H: Corilagin
R= Gal: Tercatain

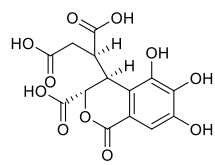


R_1= H, R_2= Gal: Gemin D
R_1=R_2= Gal: Tellimagrandin I

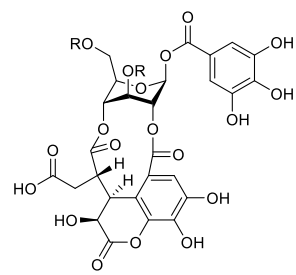
Chebulic ellagitannins



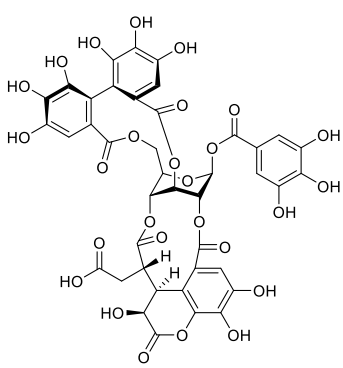
Chebulic acid



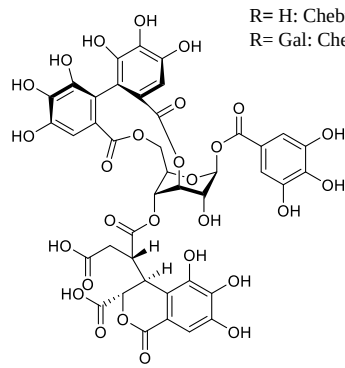
Neochebulic acid



R= H: Chebulanin
R= Gal: Chebulinic acid



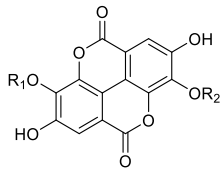
Chebulagic acid



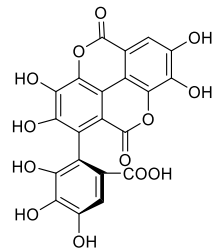
Neochebulagic acid

Figure 1. Structures of chemical constituents reported in *T. chebula*

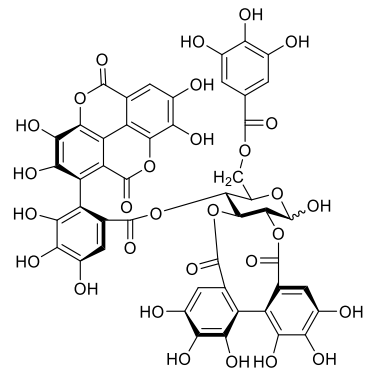
Ellagic acid and ellagic acid derivatives



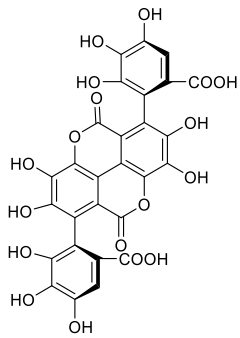
R_1=R_2=H: Ellagic acid
R_1=H, R_2=H: 3-*O*-methyl ellagic acid
R_1=R_2=Me: 3,3′-Di-*O*-methyl ellagic acid



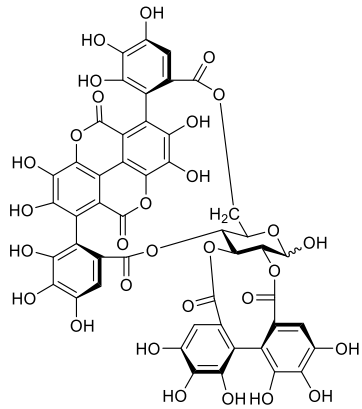
Flavogallonic acid



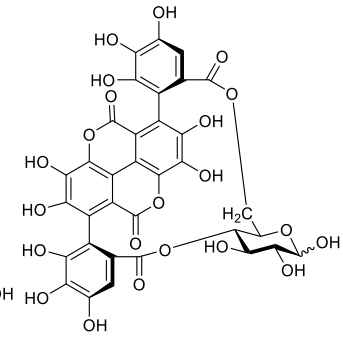
Terflavin A



Gallagic acid

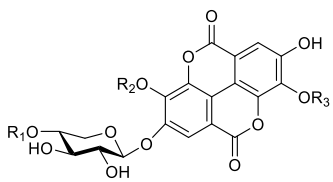


Punicalagin

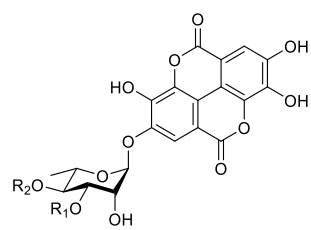


Punicalin

Ellagic acid glycosides



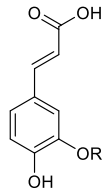
R_1=R_3=H, R_2=Me: 3′-*O*-Methyl-4-*O*-(β-D-xylopyranosyl) ellagic acid
R_1=H, R_2=Me, R_3=Gal: 3′-*O*-Methyl-4-*O*-(4″-*O*-galloyl-β-D-xylopyranosyl)ellagic acid
R_1=R_2=Me, R_3=Gal: 3,3′-*O*-Dimethyl-4-*O*-(4″-*O*-galloyl-β-D-xylopyranosyl)ellagic acid



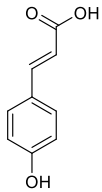
R_1=H, R_2=Gal: 4-*O*-(4″-*O*-galloyl-α-L-rhamnosyl)ellagic acid
R_1=R_2=Gal: 4-*O*-(2″,4″-di-*O*-galloyl-α-L-rhamnosyl) ellagic acid

Figure 2. Structures of chemical constituents reported in *T. chebula* (Continued)

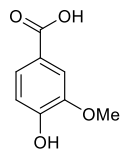
Simple phenolic acid



R=H: Caffeic acid
R=Me: Ferulic acid

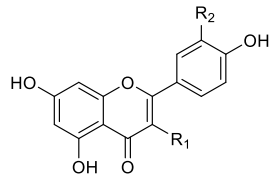


Coumaric acid



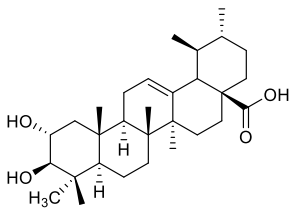
Vanillic acid

Flavonoids

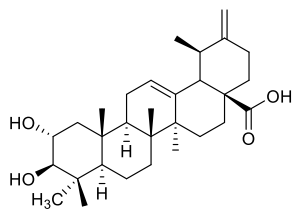


R_1=**O**H, R_2=OH: Quercetin
R_1=H, R_2=OH: Apigenin

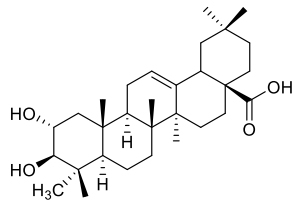
Triterpenes



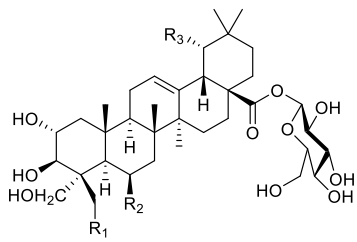
2α-Hydroxyursolic acid



2α-Hydroxymicromeric acid



Maslinic acid



R_1=R_3=OH, R_2=H: Bellericoside
R_1=R_2=H, R_3=OH: Chebuloside I
R_1=H, R_2=R_3=OH: Chebuloside II

Figure 3. Structures of chemical constituents reported in *T. chebula* (Continued)

Hydrolysable tannins of *T. chebula* species

Hydrolysable tannins are derivatives of gallic acid. Gallic acid is esterified to a core sugar (or polyol), and the galloyl moieties may be additionally esterified and oxidatively cross-linked to form more complex hydrolysable tannins (Okuda and Ito, 2011). Hydrolysable tannins of the *Terminalia* species are classified into four

groups called gllotannins, ellagitannins, chebulic ellagitannins, and ellagic acid derivatives and glycosides (Pfundstein et al., 2010).

Gallotannins

Gallotannins, are the simplest hydrolysable tannins and have core structures in which a sugar (mostly D-glucose) is esterified with gallic acid at several hydroxyl groups. Regarding partial substitution with galloyl moieties the remaining hydroxyl groups may either be substituted or unsubstituted with other diverse moieties. For instance, the anomeric hydroxyl group of the central sugar of the gallotannins may be substituted in an α or β form or unsubstituted in an α, β mixture as an ester or acetal (Khanbabaee and van Ree, 2001). The gallotannins 1,2,3,4,6-*O*-penta-galloyl-*β*-D-glucose (PGG), which are key intermediates in the biosynthesis of all hydrolysable tannins, are found in many plant families such as Combretaceae, Ericaceae, Anacardiaceae, Geraniaceae, and Aceraceae (Haddock et al., 1982; Hagenah and Gross, 1993). Gallotannins, in which the central polyol moiety is coupled to coumaroyl and cinnamoyl groups, have been rarely reported (Kashiwada et al., 1984; Saijo et al., 1989). Most of the gallotannins that are substituted with a galloyl group at the anomeric centre of their central glucose unit have the β configuration at the anomeric centre. In addition, the conformation of the central sugar moiety of gallotannins is $^4C_1$ conformation.

Ellagitannin

With numerous natural products characterized so far, ellagitannins are by far the largest group of hydrolysable tannins (Feldman and Sambandam, 1995). Oxidative coupling of at least two galloyl moieties yielding a chiral hexahydroxydiphenic

acid (HHDP) unit converts gallotannins to ellagitannins. The chirality of the HHDP unit is caused by the bulky *o*-substituents to the biaryl axis and the atropisomerism by inhibiting free rotation around the axis (Khanbabaee and van Ree, 2001). HHDP spontaneously lactonizes to ellagic acid in an aqueous solution. Most of the ellagitannins with HHDP units are linked *via* the 2,3-, 2,4-, 3,6-, or 4,6-positions of their central glucose unit (Kashiwada et al., 1992b). While the central glucose adopts a $^4C_1$ conformation in the case of a 4,6- or 2,3-HHDP linkage, a 2,4- or ,3,6-HHDP linkage always leads to the thermodynamically less stable $^1C_4$ conformation. In some plants, ellagitannins are even more elaborated through the ring opening of the central sugar converted into *C*-glycosidic ellagitannins (Konig et al., 1994). In addition to the HHDP group, other acyl units in ellagitannins include HHDP and galloyl group metabolites such as dehydrohexahydroxydiphenoyl (DHHDP), valoneoyl, gallagyl, and chebuloyl groups.

Chebulic ellagitannin

T. chebula is known as chebulic myrobalan due to the presence of a chebulic acid which is considered as a product of further ring-opening and oxidation of the DHHDP unit (Figure 4) (Walia et al., 2013; Yoshida et al., 1980). Chebulic ellagitannin refers to an ellagitannin which is bound to a chebuloyl and neochebuloyl unit. Typical chebulic ellagitannin such as chebulinic acid, chebulagic acid, neochebulagic acid, and neochebulinic acid, have been reported from the genus *Terminalia* (Pfundstein et al., 2010). The absolute configuration of chebulic acid was identified as 2*S*, 3*S*, 4*S* (Yoshida et al., 1982). Neocheulic acid has the 2*S*, 3*S*, 4*R*, configuration and is rarely found in nature (Ding et al., 2000).

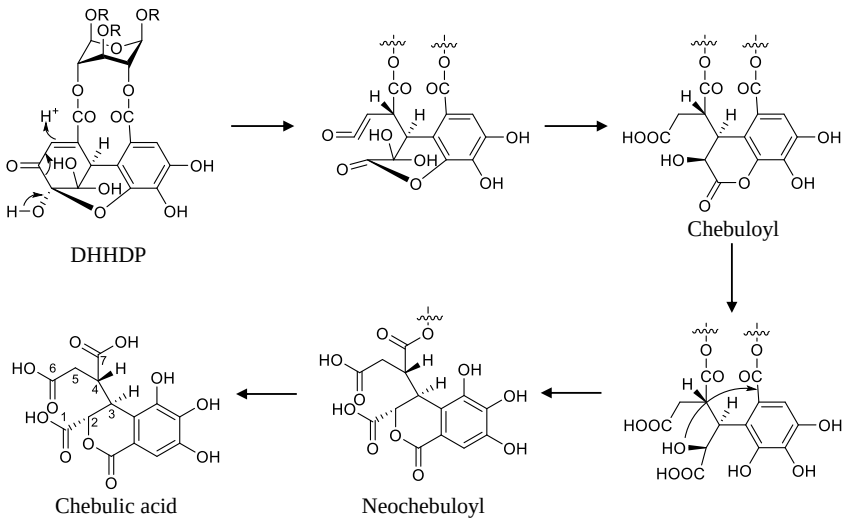


Figure 4. The putative biogenetic pathway to chebulic acid

Ellagic acid derivatives and ellagic acid glycosides

Ellagic acid, the lactonized form of HHDP, and its derivatives have also been reported from *T. chebula*. Gallagic acid is the fully lactonized form of the gallagyl unit, which is constructed from the ellagic acid nucleus which is C-C coupled with two galloyl units. Punicalin and Punicalagin are typical ellagic acid derivatives containing a gallagyl unit (Kashiwada et al., 1992a). Flavogallonic acid is also an ellagic acid derivative which is composed of an ellagic acid and a gallic acid unit linked together with a C-C bond (Marzouk et al., 2002). As a glycoside of ellagic acid, ellagic acid xylosides and their galloyl derivatives have been mainly reported from genus *Terminalia* (Fahmy et al., 2015). The glycosidic linkage of the xyloside was identified to be at position C-4 of the ellagic acid. In addition to the xylosides, ellagic acid rhamnosides and their galloyl derivatives have been reported from *T. chebula* a few years ago (Pfundstein et al., 2010).

HPLC-based activity profiling

For thousands of years, medicinal plants have been used as therapeutic agents for humans. The bioactive constituents in medicinal plants have also been investigated for centuries (Cragg and Newman, 2013). Through these investigations, from morphine to taxol, many natural products and their derivatives have been developed for clinical use to treat human diseases in almost all therapeutic areas and are considered a continuing source of novel drug leads (Newman et al., 2000).

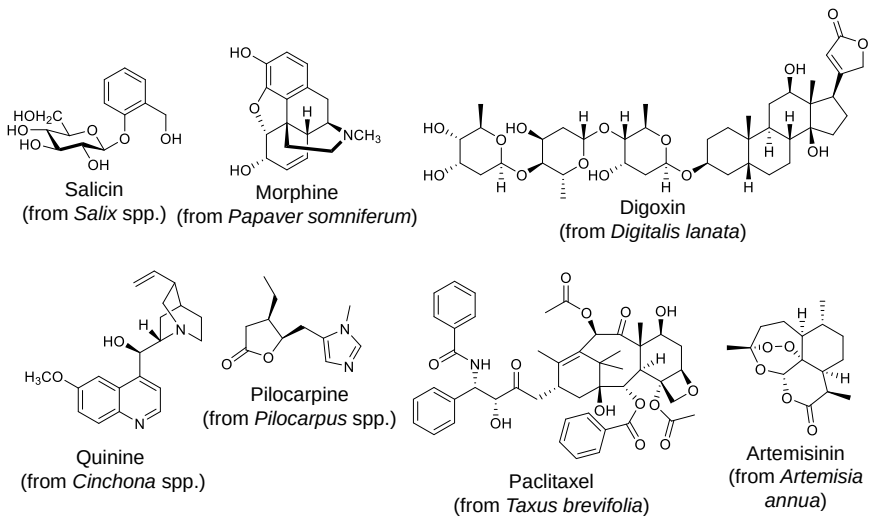


Figure 5. Drugs from natural products

Bioactivity-guided isolation is the classical approach of natural product based drug discovery. A bioactive extract is fractionated and purified under the guidance of a specific bioassay to isolate biologically active constituents. This approach is a time-consuming and costly process (Hostettmann et al., 2005). All fractions from each isolation step need to be tested for their bioactivity, and structure elucidation of the pure compound is performed at the end of the process. Furthermore, already known compounds are often re-isolated. Many approaches to effective isolate

bioactive natural products have been developed.

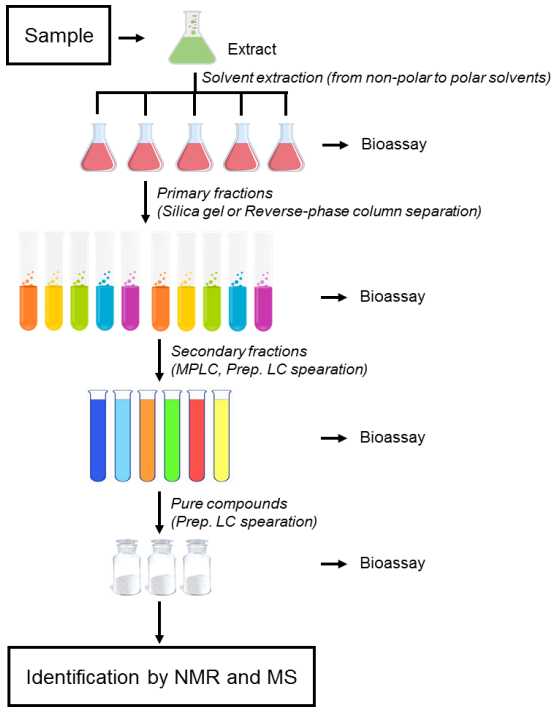
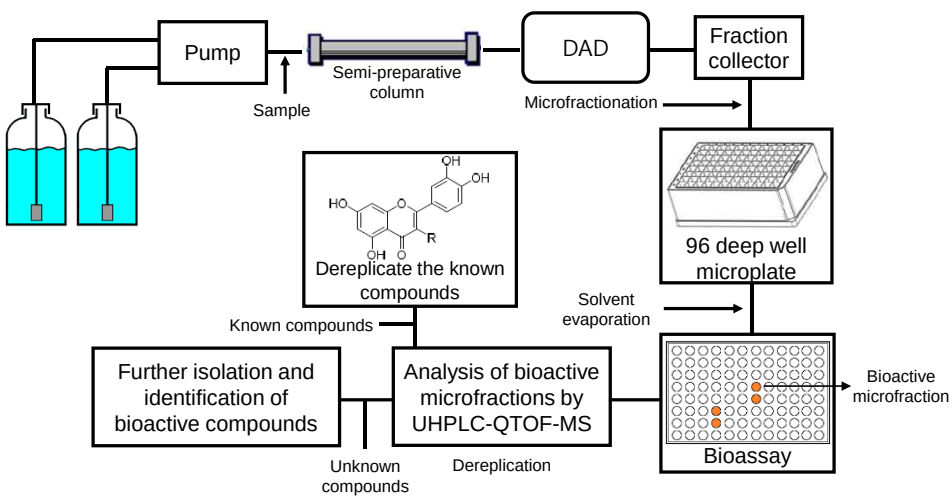


Figure 6. The general concept for bioactivity-guided isolation

The remarkable improvement in the performance of HPLC columns and the availability of HPLC hyphenated techniques have enabled acquisition of much structural information on the constituents contained in a bioactive extract of a fraction in a short time (Seger et al., 2013). The combination of HPLC hyphenated techniques with bioassays have enabled the correlation of structural information of discrete peaks with their bioactivity and therefore, tracking bioactive constituents in the bioactive extract or fraction. Three different approaches, on-line post-column bioassays, at-line settings, and off-line activity profiling, have been developed. However, among the above three approaches, the off-line approach, which is commonly referred to as HPLC-based activity profiling, is considered the most versatile approach with potential for broad applications in drug discovery from

natural products (Potterat and Hamburger, 2013). In HPLC-based activity profiling, milligram amounts of active extracts or fractions are separated by analytical or semi-preparative HPLC. UV spectra and HRMS data are measured and time-based fractions are collected in 96 deep-well microplates. Separated fractions are dried by a centrifugal evaporator or freeze dryer, re-dissolved in a small amount of an appropriate solvent, and tested for their bioactivity. The activity profile of each fraction and the chromatogram are matched manually to identify active peaks. In addition, structural information such as the UV spectrum and high resolution MS of the active peaks help to determine whether the peaks are likely known or new through database searches (*i.e.*, dereplication). Furthermore, a straightforward peak-based targeted isolation of bioactive constituents can be performed without a bioassay for each isolation step (Scheme 1). In this thesis, the same HPLC-based activity profiling approach was used for tracking the α-glucosidase inhibitor in *T. chebula* fruits.



Scheme 1. Schematic representation of HPLC-based activity profiling with dereplication

α-Glucosidase and its inhibitory agents from natural products

Type 2 diabetes mellitus (T2DM), known as non-insulin-independent diabetes, is a chronic metabolic disorder due to insulin-resistance in tissues (Matthews et al., 1985). Insulin resistance results in elevated blood glucose levels which can damage many of the organs (Guo, 2014). One of the therapeutic approaches for treating T2DM is to alleviate postprandial hyperglycemia. This is done by suppressing glucose absorption from the gut by inhibiting the intestinal carbohydrate digesting enzymes such as α-glucosidase (EC number 3.2.1.20) (Kim et al., 2000). α-Glucosidase, which is a membrane-bound enzyme located in the epithelium of the small intestine, catalyzes the hydrolytic reaction to liberate glucose from the non-reducing end of the oligosaccharide (Hirsh et al., 1997). Acarbose is a clinical α-glucosidase inhibitor which has been used as a first-line treatment for diabetic patients with postprandial hyperglycemia (Sivasothy et al., 2016). Acarbose is efficient in attenuating the rapid increase in the blood glucose levels of patients; however its continuous use may cause side effects such as flatulence and diarrhea (Chiasson et al., 2002). Thus, much effort has been given to searching for effective and safe α-glucosidase inhibitors from natural materials to develop bioactive functional foods to treat T2DM (Table 2).

Table 2. Potential natural products for inhibition against α-glucosidase

Compound	Source	Type of compound	Reference
3β-acetoxy-16β-hydroxybetulinic acid	*Fagara tessmannii*	Triterpene	(Mbaze et al., 2007)
Segetalic acid glycosides	*Gypsophila oldhamiana*	Triterpene saponin	(Luo et al., 2008)
(25*S*)-5α-furastan-3β,22,26-triol	*Tribulus longipetalus*	Steroid	(Naveed et al., 2014)
D-Galactopyranosyl harpagoside	*Scrophularia ningpoensis*	Iridoid	(Hua et al., 2014)
Deoxynojirimycin	*Lobelia sessilifolia*	Aminoglycoside	(Ikeda et al., 2000)
Piperumbellactams A	*Piper umbellatum*	Alkaloid	(Tabopda et al., 2008)
Bisgerayafoline D	*Murraya koenigii*	Alkaloid	(Uvarani et al., 2014)
Rhapontigenin	*Rheum emodi*	Stilbene	(Babu et al., 2004)
Quercetin	*Matricaria recutita*	Flavonoid	(Kato et al., 2008)
Cudratricusxanthone F	*Cudrania tricuspidata*	Xanthone	(Seo et al., 2007)

II. Materials and Methods

1. Materials

1.1. Plant material

T. chebula fruits, which were collected from the Yunnan province, China, were provided by Chong Kun Dang Pharmaceutical Corp. (Seoul, Republic of Korea). A voucher specimen (SUPH-2015-01) has been deposited at the Herbarium in the Medicinal Plant Garden, Seoul National University.

1.2. Reagents for isolation and purification

First grade solvents (Daejung, Si-heung, Republic of Korea) were used for extraction, fractionation, and isolation. HPLC-grade solvents were purchased from J. T. Baker Chemical Corp. (Center valley, PA, USA). Column chromatography (CC) was performed with the Diaion HP-20 (Mitsubishi Chemical Industries Ltd., Tokyo, Japan) and Sephadex LH-20 (25-100 μm, GE Healthcare, IL, USA). Thin-layer chromatography (TLC) was performed the pre-coated normal silica gel 60 F254 and RP-C18 F-254S (Merck). Methanol-d_4 (CD_3OD), dimethyl sulfoxide-d_6 (DMSO-d_6), acetone-d_6, pyridine-d_5, and deuterium oxide (D_2O) were used as NMR solvents (Merck, Darmstadt, Germany). All other reagents were purchased from Sigma-Aldrich (MO, USA).

1.3. Reagents for *in vitro* enzyme assay

Baker's yeast α-glucosidase (Cat. No. G5003), rat intestinal acetone powder (Cat. No. I1630), porcine pancreatic α-amylase (Cat. No. A3176), *p*-nitrophenyl-*α*-D-glucopyranoside (pNGP, Cat. No. 71768), potato starch (Cat. No. S2004), iodine solution (Cat. No. 15-0630), hydrochloric acid (Cat. No. 320331), *p*-nitrophenylphosphate (pNPP, Cat. No. P4744), dithiothreitol (Cat. No. 43819), sodium chloride (NaCl, Cat. No. S7653), ethylenediaminetetraacetic acid (EDTA, Cat. No. EDS), sodium phosphate monobasic (NaH_2PO_4, Cat. No. S8282), sodium phosphate dibasic (Na_2HPO_4, Cat. No. S9390), citric acid (Cat. No. C1909), sodium citrate dihydrate (Cat. No. C3434), and acarbose (Cat. No. PHR1253) were purchased from Sigma-aldrich (St. Louis, Mo. USA). PTP-1B (human recombinant) was purchased from Enzo life sciences (Farmingdale, NY, USA)

1.4. Equipments

Analytical balance: Mettler Ae 50, Switzerland

Analytical HPLC: Dionex P680 equipped with a Dionex UVD340U detector

Centrifuge: Effendorff centrifuge 5810, Germany

Centrifugal evaporator:

Drying oven: CO-2D-1S, Wooju Sci., Republic of Korea

Evaporator: EYELA NE, Japan

Fraction collector: Waters fraction collector III, Waters, UK

FT-IR: Jasco FT/IR-4200 spectrophotometer, USA

Freeze-dryer: DURA-DRY, Fts system Inc., USA

HPLC column: YMC-Triart C18 (4.6×250 mm, 10.0×250 mm, and 20.0×250 mm 5 μm), YMC Co., Ltd., Japan

Microplate reader: SpectraMax M5, Molecular Devices, CA, USA

MPLC: MPLC-Reveleris system, Grace, USA

MPLC column: Reveleris flash cartridges (C18, 330 g), Grace, USA

NMR: JEOL LA 300, GSX 400, JEOL, Japan

Bruker Avance-600, 800, Bruker, Germany

Polarimeter: Jasco P-2000 polarimeter, Jasco, USA

Ultrasonicator: Branson 5210, UK

UV/CD spectrometer: Chirascan-plus spectropolarimeter, Applied Photophysics Ltd, UK

Semipreparative HPLC: Waters Delta Prep 400 system equipped with a Waters 2489 UV/Vis detector, Waters, UK

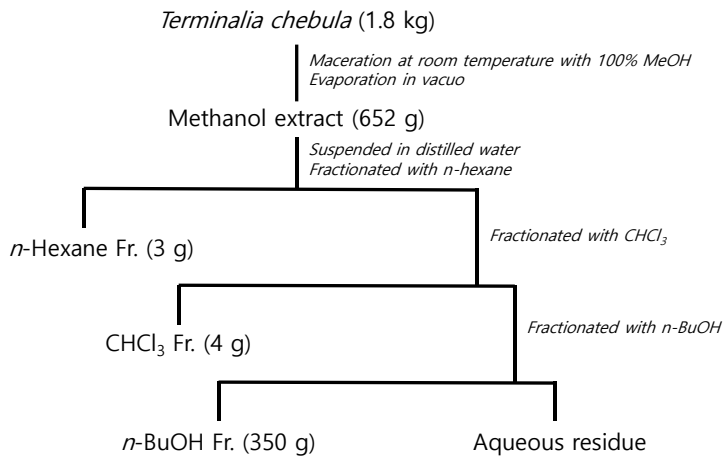
UHPLC column: Waters Acquity BEH C18 column (1.7 μm, 2.1 × 100 mm), Waters, UK

UHPLC-QTOF-MS: Waters Xevo G2 qTOF mass spectrometer with Acquity UPLC system, Waters, UK

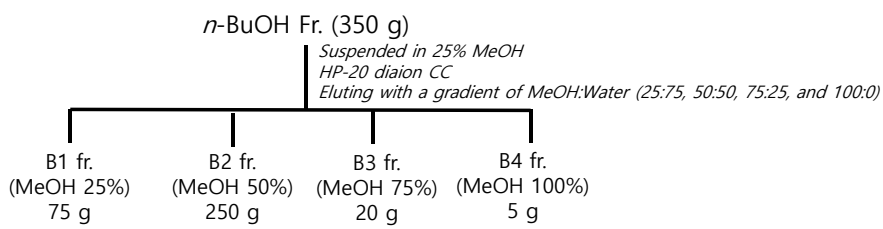
2. Methods

2.1. Extraction and fractionation of *T. chebula*

Air-dried *T. chebula* fruits (1.8 kg) were macerated with MeOH (8 L × 2) at room temperature (12 h × 2). The total extract (652 g) which was evaporated *in vacuo* was diluted with distilled water (4 L) and successively fractionated into *n*-hexane fraction (3 g), $CHCl_3$ fraction (4 g) and n-BuOH fraction (350 g), respectively. The n-BuOH fraction was subjected to HP-20 resin (1.6 kg) column chromatography eluting with a gradient of MeOH:H_2O (25:75, 50:50, 75:25, 100:0, each 8 L) to give four sub-fractions (B1–B4).



Scheme 2. Extraction and fractionation of *T. chebula* fruits



Scheme 3. Further fractionation of *n*-BuOH fraction of *T. chebula* fruits

2.2. HPLC-based activity profiling of *T. chebula*

HPLC-based activity profiling was performed by microfractionation with a semipreparative YMC-Triart C18 column (10×250 mm, 5 μm). The solvent system used a mixture of aqueous 0.1% formic acid (A) and MeCN (B), and the following gradient was applied: 16% to 25% B in 30 min (B2 fraction) and 17% to 47% B in 30 min (B3 fraction). The flow rate was 3.4 mL/min and the column termperature was maintained at 30℃. The subfraction B2 and B3 were dissolved in MeOH (10 mg/ml) and the injection volume was 50 μL (0.5 mg of extract was injected). The UV absorbance was measured at 280 nm. The microfractions were collected every 0.5 min from 8 min to 30 min in a 96 deep-well plate (1.7 mL per well) by fraction collector. After collection, the plate was dried by centrifugal evaporator. The contents of plate were dissolved in 150 μL of 50 mM PBS buffer (pH 7.0) and used for the determination of the Baker's yeast α-glucosidase inhibition.

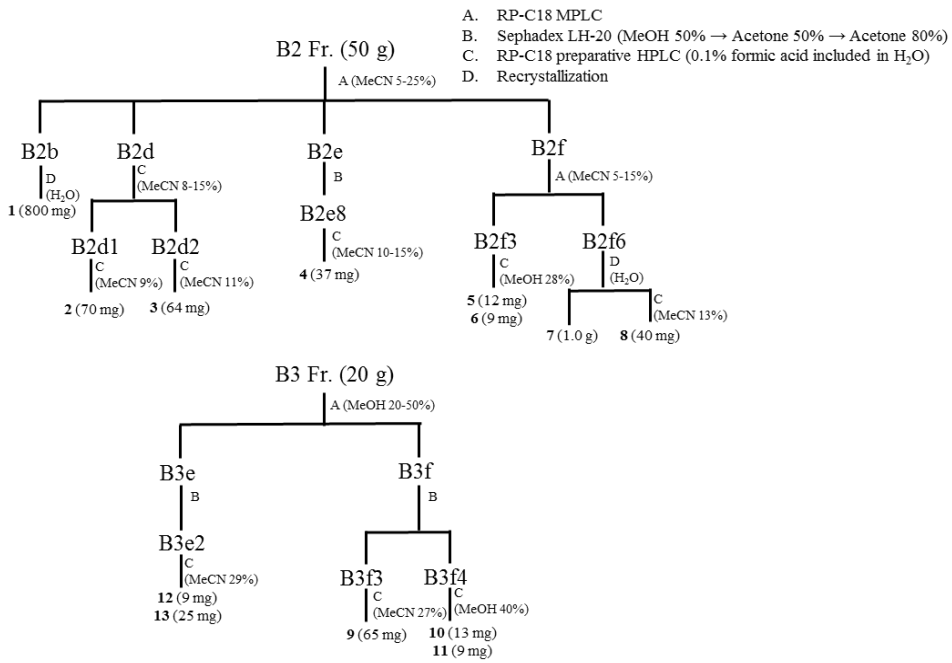
The measurement of the Baker's yeast α-glucosidase inhibitory activity was performed as described by Jabeen et al. with some modification (Jabeen et al., 2013). Briefly, 50 μL of 180 mU/mL α-glucosidase in PBS buffer and 100 μL of the each microfractions in PBS buffer were mixed and incubated at 36℃ for 10 min in 96-well plate. Then, 50 μL of 1.2 mM *p*-nitrophenyl-α-D-glucopyranose (pNGP) was added to initiate the reaction. The reaction mixture was incubated at 36℃ and monitored using a microplate reader at a wavelength of 405 nm, which indicated the formation of *p*-nitrophenol. Absorbance of mixture was measured every 2.5 minute up to 5 minute. The reaction mixture with the test solution replaced by equivalent PBS buffer was used as control, and acarbose was used as a positive control. Inhibition of α-glucosidase was calculated as following equation: % inhibition = [1-(rate of sample reaction/rate of control reaction)]×100. All the samples were performed in triplicate.

Dereplication of each active microfractions was performed by using a UHPLC-QTOF-MS. The mobile phases were aqueous 0.1% formic acid (A) and 100% MeCN (B), with the following gradient: 3-20% B (0-14 min), 20-90% B (14-20 min), 100% B (20-23 min), and 3% B (23-26 min). The flow rate was set at 300 μL/min. The ESI conditions were set as follows: negative (-) ion mode; capillary voltage, 2.5 kV; cone voltage, 40 V; source temperature, 120℃; desolvation temperature, 350℃; cone gas flow, 50 L/h; and desolvation gas flow, 600 L/h. The energy for collision-induced dissociation (CID) was set to 3 V for the precursor ion, and the MS fragment information obtained using a collision energy ramp from 20 eV to 40 eV in MSe mode (Plumb et al., 2006). To ensure the mass accuracy of the MS acquisition, leucine encephalin (Tyr-Gly-Gly-Phe-Leu, m/z 554.2615 in negative mode) was used as the lock mass at a concentration of 200 ng/mL and a flow rate of 3 μL/min. Calibration of the MS system in the range from m/z 100 to 1500 was performed using a sodium formate solution (5 μM/mL). Molecular formula which was derived from high resolution MS data, UV spectrum, and MS fragment of each active microfractions were used for a narrowed search in Chemical Abstracts Service (CAS) Registry via SciFinder database.

2.3. Targeted Isolation of Baker's yeast α-glucosidase inhibitory compounds from B2 and B3 fraction of *T. chebula*

B2 fraction (50 g) was subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (5:95 to 30:70) to yield 6 subfractions (B2a-B2f). Fraction B2b was recrystallized using water to afford **1** (800 mg). Fraction B2d was

separated by semi-preparative HPLC, eluted with a MeCN/ H_2O with 0.1% formic acid (8:92 to 15:85) to yield 2 subfractions (B2d1 and B2d2). Compound **2** (70 mg) and **3** (64 mg) were obtained from B2d1 and B2d2 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 9:91 and 11:89), respectively. Fraction B2e was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) and purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 10:90 to 15:85) to give compound **4** (37 mg). Fraction B2f was subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (5:95 to 15:85) to yield seven subfractions (B2f1-B2f7). Fraction B2f3 was separated on semi-preparative HPLC column (MeOH/H_2O with 0.1% formic acid 28:72) to afford compound **5** (12 mg) and **6** (9 mg). Fraction B2f6 was recrystallized using water to afford compound **7** (1.0 g) and its filtrate yielded compound **8** (40 mg) after separation on semi-preparative HPLC column (MeCN/H_2O with 0.1% formic acid 13:87). B3 fraction was subjected to MPLC eluted with a MeOH/H_2O with 0.1% formic acid gradient (20:80 to 50:50) to yield seven subfractions (B3a-B3g). Fraction B3e was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield three subfractions (B3e1-B3e3). Fraction B3e2 was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 29:71) to give compound **12** (9 mg) and **13** (25 mg). Fraction B3f was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield five subfractions (B3f1-B3f5). Compound **9** (65 mg) was obtained from fraction B3f3 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 27:73). Compound **10** (13 mg) and **11** (9 mg) were obtained from fraction B3f4 through semi-preparative HPLC (MeOH/H_2O with 0.1% formic acid 40:60)



Scheme 4. Targeted isolation of compounds from *T. chebula* fruits

1,3,6-Tri-*O*-galloyl-*β*-D-glucose (**1**)

White amorphous powder

$C_{27}H_{24}O_{18}$

$[\alpha]_D^{20}$= 30.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 278 (4.5) nm

IR ν_{max}: 3388, 1700, 1616, 1537, 1454, 1342, 1213 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 635.0892 $[M - H]^-$ (calcd for $C_{27}H_{24}O_{18}$, 635.0884)

^{1}H (300 MHz, CD_3OD) NMR: δ 7.16 (2H, *s*, Gal-H-2″, 6″), 7.13 (2H, *s*, Gal-H-2′, 6′), 7.09 (2H, *s*, Gal-H-2‴, 6‴), 5.83 (1H, *d*, *J* = 8.2 Hz, H-1), 5.27 (1H, *t*, *J* = 9.3

Hz, H-3), 4.58 (1H, *dd*, *J* = 1.6 12.1 Hz, H-6a), 4.45 (1H, *dd*, *J* = 4.6, 12.1 Hz, H-6b), 3.88 (1H, *m*, H-5), 3.78 (1H, *dd*, *J* = 5.5, 9.5 H-4) 3.77 (1H, *dd*, *J* = 4.5, 9.5 Hz, H-2)

^{13}C (75 MHz, CD_3OD) NMR: δ 168.1, 168.0, 166.7 (Gal-C-7′, C-7″, C-7‴), 146.5, 146.4, 146.4 (Gal-C-3′, 5′, C-3″, 5″, C-3‴, 5‴), 140.4, 139.8, 139.7 (Gal-C-4′, C-4″, C-4‴), 121.5, 121.2, 120.4 (Gal-C-1′, C-1″, C-1‴), 110.5, 110.3, 110.1 (Gal-C-2′, 6′, C-2″, 6″, C-2‴, 6‴), 95.8 (C-1), 78.9 (C-3), 76.4 (C-5), 72.6 (C-2), 69.7 (C-4), 64.2 (C-6)

Chebulagic acid (**2**)

Yellowish amorphous powder

$C_{41}H_{30}O_{27}$

$[\alpha]_D^{20}$= -37.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 278 (4.5) nm

IR ν_{max}: 3485, 1711, 1617, 1513, 1457, 1340, 1211 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 953.0865 $[M - H]^-$ (calcd for $C_{41}H_{29}O_{27}$, 953.0896)

^{1}H (400 MHz, CD_3OD) NMR: δ 7.48 (1H, *s*, Cheb-H-2″), 7.07 (2H, *s*, Gal-H-2′, 6′), 6.84 (1H, *s*, HHDP-H-5′), 6.63 (1H, *s*, HHDP-H-5″), 6.51 (1H, *s*, H-1), 5.83 (1H, *s*, H-3), 5.39 (1H, *s*, H-2), 5.23 (1H, *d*, *J* = 3.6 Hz, H-4), 5.05 (1H, *dd*, *J* = 1.2, 7.3 Hz, Cheb-H-3′), 4.90 (1H, *m*, H-6a), 4.82 (1H, *m*, H-5), 4.80 (1H, *m*, Cheb-H-2′), 4.37 (1H, *dd*, *J* = 7.6, 10.4 Hz, H-6b), 3.80 (1H, *ddd*, *J* = 1.1, 3.6, 11.7 Hz, Cheb-H-4′), 2.14 (2H, *m*, Cheb-H-5′)

^{13}C (100 MHz, CD_3OD) NMR: δ 175.0 (Cheb-C-6′), 174.4 (Cheb-C-7′), 170.7

(Cheb-C-1′), 170.1 (HHDP-C-7″), 167.5 (HHDP-C-7′), 166.4 (Gal-H-7′), 166.2 (Cheb-C-7″), 147.4 (Cheb-C-3″), 146.5 (Gal-C-3′, 5′), 146.1 (HHDP-C-4″), 145.6 (HHDP-C-4′), 145.5 (HHDP-C-2′), 145.3 (HHDP-C-2″), 141.4 (Cheb-C-5″), 140.8 (Gal-C-4′), 140.4 (Cheb-C-4″), 138.6 (HHDP-C-3′), 137.5 (HHDP-C-3″), 125.6 (HHDP-C-6″), 124.5 (HHDP-C-6′), 120.1 (Gal-C-1′), 119.0 (Cheb-C-1″), 117.6 (Cheb-C-2″), 117.6 (HHDP-C-1′) 116.2 (HHDP-C-1″), 116.0 (Cheb-C-6″), 110.9 (Gal-C-2′, 6′), 110.4 (HHDP-C-5′), 108.2 (HHDP-C-5″), 92.5 (C-1), 74.2 (C-5), 71.1 (C-2), 67.0 (Cheb-C-2′), 66.8 (C-4) , 64.7 (C-6), 62.4 (C-3), 41.7 (Cheb-C-3′), 40.0 (Cheb-C-4′), 30.5(Cheb-C-5′)

6′-*O*-Methyl neochebulagate (**3**)

Tan amorphous powder

$C_{42}H_{34}O_{28}$

$[\alpha]_D^{20}$= -52.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 221 (4.9), 279 (4.5) nm

CD (MeOH): 214 (Δε -4.77), 235 (Δε 5.07), 261 (Δε -0.26), 291 (Δε 1.93)

IR ν_{max}: 3387, 2948, 1725, 1516, 1365, 1203, 1024, 677 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 985.1151 $[M - H]^-$ (calcd for $C_{42}H_{33}O_{28}$ 985.1158)

^{1}H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: 7.10 (2H, *s*, Gal-H-2′, 6′), 6.83 (1H, *s*, HHDP-H-5′), 6.70 (1H, *s*, HHDP-H-5″), 6.16 (1H, *d*, *J* = 4.0 Hz, H-1), 5.67 (1H, *d*, *J* = 3.3 Hz, H-4), 5.42 (1H, *d*, *J* = 0.7 Hz, Neocheb-H-2′), 4.90 (1H, *d*, *J* = 3.1 Hz, H-3), 4.59 (1H, *m*, H-6a), 4.56 (1H, *m*, H-5), 4.29 (1H, *dd*, *J* = 6.9, 10.5 Hz, H-6b), 4.20 (1H, br *d*, *J* = 4.0 Hz, H-2), 3.98 (1H, *d*, *J* = 9.1 Hz, Neocheb-H-3′), 3.52 (3H,

s, OMe), 3.27 (1H, *td*, *J* = 9.1, 5.1 Hz, Neocheb-H-4′), 2.83 (1H, *dd*, *J* = 9.0, 17.3 Hz, Neocheb-H-5′a), 2.51 (1H, *dd*, *J* = 5.1, 17.3 Hz, Neocheb-H-5′b)

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: 172.8 (Neocheb-C-6′), 172.6 (Neocheb-C-7′), 171.5 (Neocheb-C-1′), 168.8 (HHDP-C-7″), 167.1 (HHDP-C-7′), 165.8 (Gal-C-7′), 164.9 (Neocheb-C-7″), 146.1 (Neocheb-C-3″), 145.9 (Gal-C-3′, 5′), 145.2 (HHDP-C-4″), 144.8 (HHDP-C-4′), 144.8 (HHDP-C-2′), 144.8 (HHDP-C-2″), 143.3 (Neocheb-C-5″), 139.6 (Gal-C-4′), 139.3 (Neocheb-C-4″), 137.0 (HHDP-C-3′), 136.4 (HHDP-C-3″), 125.1 (HHDP-C-6″), 125.0 (HHDP-C-6′), 120.2 (Gal-C-1′), 116.8 (Neocheb-C-6″), 116.5 (HHDP-C-1′), 116.4 (Neocheb-C-1″), 115.5 (HHDP-C-1″), 110.3 (Gal-C-2′, 6′), 109.7 (HHDP-C-5′), 108.9 (Neocheb-C-2″), 108.2 (HHDP-C-5″), 94.7 (C-1), 78.1 (Neocheb-C-2′), 74.2 (C-5), 71.8 (C-3), 69.9 (C-2), 65.1 (C-4), 64.5 (C-6), 52.2 (OMe), 44.1 (Neocheb-C-4′), 36.0 (Neocheb-C-3′), 34.2 (Neocheb-C-5′)

1′-*O*-Methyl neochebulinate (**4**)

Yellowish amorphous powder

$C_{42}H_{36}O_{28}$

$[\alpha]_D^{20}$= 18.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (5.0), 281 (4.6) nm

IR ν_{max}: 3390, 1710, 1617, 1536, 1456, 1342, 1209 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 987.1348 $[M - H]^-$ (calcd for $C_{42}H_{35}O_{28}$ 987.1315)

^{1}H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 7.16 (2H, *s*, Gal-H-2′, 6′) 7.15 (2H, *s*, Gal-H-2″, 6″), 7.15 (2H, *s*, Gal-H-2‴, 6‴), 7.05 (1H, *s*, Neocheb-H-2″), 6.02 (1H,

d, *J* = 8.4 Hz), 5.66 (1H, *t*, *J* = 9.5 Hz, H-3), 5.53 (1H, *t*, *J* = 9.7 Hz, H-4), 4.80 (1H, *d*, *J* = 1.2 Hz, Neocheb-H-2′), 4.62 (1H, br *d*, *J* = 12.4 Hz, H-6a), 4.41 (1H, *dd*, *J* = 4.4, 12.5 Hz, H-6b), 4.34 (1H, *ddd*, *J* = 2.0, 4.4, 10.2 Hz, H-5), 3.99 (1H, *m*, H-2), 3.73 (1H, *dd*, *J* = 1.6, 11.2 Hz, Neocheb-3′), 3.28 (3H, *s*, OMe), 3.15 (1H, *ddd*, *J* = 4.3, 9.9, 11.2 Hz, Neocheb-H-4″), 2.86 (1H, *dd*, *J* = 10.0, 17.7 Hz, Neocheb-5′a), 2.37 (1H, *dd*, *J* = 4.3, 17.7 Hz, Neocheb-5′b)

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 174.0 (Neocheb-C-7′), 173.2 (Neocheb-C-6′), 169.7 (Neocheb-C-1′), 166.5 (Gal-C-7‴), 166.3 (Gal-C-7″), 165.4 (Gal-C-7′), 163.7 (Neocheb-C-7″), 146.1 (Gal-C-7‴), 146.1 (Gal-C-3″, 5″) 146.0 (Gal-C-3′, 5′), 146.0 (Neocheb-C-3″), 143.5 (Neocheb-C-5″), 139.6 (Gal-C-4′), 139.1 (Gal-4″), 139.1 (Neocheb-C-4″), 138.9 (Gal-4‴), 121.7 (Gal-C-1‴), 121.1 (Gal-C-1″), 120.5 (Gal-C-1′), 116.7 (Neocheb-C-1″), 110.4 (Gal-C-2‴, 6‴), 110.4 (Gal-C-2″, 6″) 110.1 (Gal-C-2′, 6′), 109.3 (Neocheb-C-2″), 95.2 (C-1), 77.4 (Neocheb-C-2′), 75.9 (C-3), 73.5 (C-5), 72.7 (C-2), 70.0 (C-4), 62.8 (C-6), 52.8 (OMe), 45.1 (Neocheb-C-4′) , 36.6 (Neocheb-C-3′), 34.7 (Neocheb-C-5′)

1,2,3,6-Tetra-*O*-galloyl-*β*-D-glucose (**5**)

White amorphous powder

$C_{34}H_{28}O_{22}$

$[\alpha]_D^{20}$= 19.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 280 (4.5) nm

IR ν_{max}: 3363, 1701, 1617, 1537, 1456, 1343, 1210 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 787.1005 $[M - H]^-$ (calcd for $C_{34}H_{27}O_{22}$, 787.0994)

^{1}H (400 MHz, acetone-d_6) NMR: δ 7.15, 7.07, 7.05, 6.97 (each 2H, *s*, Gal-H-2′, 6′, Gal-H-2″, 6″, Gal-H-2‴, 6‴, Gal-H-2⁗, 6⁗), 6.16 (1H, *d*, *J* = 8.4 Hz, H-1), 5.66 (1H, *t*, *J* = 9.9 Hz, H-3), 5.45 (1H, *dd*, *J* = 8.4 9.9 Hz, H-2), 4.56 (2H, *m*, H-6), 4.12 (1H, *m*, H-5), 4.11 (1H, *m*, H-4)

^{13}C (100 MHz, acetone-d_6) NMR: δ 166.6, 166.1, 165.7, 165.0 (Gal-C-7′, Gal-C-7″, Gal-C-7‴, Gal-C-7⁗), 146.1, 146.0, 145.9 (2C) (Gal-C-3′, 5′, Gal-C-3″, 5″, Gal-C-3‴, 5‴, Gal-C-3⁗, 5⁗), 139.6, 139.10, 138.9, 138.9 (Gal-C-4′, Gal-C-4″, Gal-C-4‴, Gal-C-4⁗) 121.6, 121.3, 120.7, 120.1 (Gal-C-1′, Gal-C-1″, Gal-C-1‴, Gal-C-1⁗), 110.3, 110.1, 110.1, 110.0 (Gal-C-2′, 6′, Gal-C-2″, 6″, Gal-C-2‴, 6‴, Gal-C-2⁗, 6⁗), 93.5 (C-1), 76.0 (C-5), 75.8 (C-2), 71.7 (C-3), 69.3 (C-4), 63.6 (C-6)

1,3,4,6-Tetra-*O*-galloyl-*β*-D-glucose (**6**)

White amorphous powder

$C_{34}H_{28}O_{22}$

$[\alpha]_D^{20}$= 19.8 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 280 (4.6) nm

IR ν_{max}: 3373, 1701, 1617, 1537, 1456, 1344, 1212 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 787.1002 [M – H]$^-$ (calcd for $C_{34}H_{27}O_{22}$, 787.0994)

^{1}H (400 MHz, acetone-d_6) NMR: δ 7.20, 7.15, 7.06, 7.04 (each 2H, *s*, Gal-H-2′, 6′, Gal-H-2″, 6″, Gal-H-2‴, 6‴, Gal-H-2⁗, 6⁗), 6.04 (1H, *d*, *J* = 8.4 Hz, H-1), 5.69 (1H, *t*, *J* = 9.5 Hz, H-3), 5.49 (1H, *t*, *J* = 9.7 Hz, H-4), 4.48 (1H, br *d*, *J* = 11.1 Hz, H-6a), 4.37 (1H, *m*, H-5), 4.33 (1H, *m*, H-6b), 4.04 (1H, *t*, *J* = 8.8 Hz, H-2)

^{13}C (100 MHz, acetone-d_6) NMR: δ 166.4, 166.1, 165.7, 165.2 (Gal-C-7′, Gal-C-7″,

Gal-C-7‴, Gal-C-7⁗), 146.1, 145.9, 145.9, 145.8 (Gal-C-3′, 5′, Gal-C-3″, 5″, Gal-C-3‴, 5‴, Gal-C-3⁗, 5⁗), 139.5, 139.2, 138.9, 138.8 (Gal-C-4′, Gal-C-4″, Gal-C-4‴, Gal-C-4⁗) 121.5, 121.4, 120.7, 120.6 (Gal-C-1′, Gal-C-1″, Gal-C-1‴, Gal-C-1⁗), 110.4, 110.2, 110.1 (2C) (Gal-C-2′, 6′, Gal-C-2″, 6″, Gal-C-2‴, 6‴, Gal-C-2⁗, 6⁗), 95.4 (C-1), 75.7 (C-5), 73.8 (C-2), 71.7 (C-3), 69.3 (C-4), 63.6 (C-6)

Chebulinic acid (**7**)

White amorphous powder

$C_{41}H_{32}O_{27}$

$[\alpha]_D^{20}$= 40.9 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 280 (4.0) nm

IR ν_{max}: 3473, 1701, 1513, 1341 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 955.1057 $[M - H]^-$ (calcd for $C_{41}H_{31}O_{27}$, 955.1053)

^{1}H (400 MHz, CD_3OD) NMR: δ 7.51 (1H, *s*, Cheb-H-2″), 7.17 (2H, *s*, Gal-H-2″, 6″), 7.10 (2H, *s*, Gal-H-2′, 6′), 6.98 (2H, *s*, Gal-2‴, 6‴), 6.49 (1H, *d*, *J* = 2.8 Hz, H-1), 6.24 (1H, br *s*, H-3), 5.43 (1H, *d*, *J* = 1.2 Hz, H-2), 5.09 (1H, *dd*, *J* = 1.4, 7.2 Hz, Cheb-H-3″), 5.04 (1H, *d*, *J* = 3.4 Hz, H-4), 4.82 (1H, *m*, Cheb-H-2′), 4.81 (1H, *m*, H-6a), 4.72 (1H, *m*, H-5) 4.62 (1H, *dd*, *J* = 6.3 10.6 Hz, H-6b), 3.87 (1H, *ddd*, *J* = 0.9, 4.7, 10.5 Hz, Cheb-H-4′), 2.24 (2H, *m*, Cheb-H-5′)

^{13}C (100 MHz, CD_3OD) NMR: δ 175.1 (Cheb-C-6′), 174.6 (Cheb-C-7′), 170.7 (Cheb-C-1′), 167.9 (Gal-C-7‴), 166.3 (Gal-C-7′), 166.1 (Gal-C-7″), 166.1 (Cheb-C-7″), 147.4 (Cheb-C-3″), 146.7 (Gal-C-3″, 5″), 146.6 (Gal-C-3′, 5′), 146.4 (Gal-C-3‴, 5‴), 141.5 (Cheb-C-5″), 140.8 (Gal-C-4″), 140.8 (Gal-C-4′), 140.5 (Cheb-

C-4″), 140.1 (Gal-C-4‴), 120.7 (Gal-C-1‴), 120.0 (Gal-C-1′), 119.8 (Gal-C-1″), 119.1 (Cheb-C-1″), 117.6 (Cheb-C-2″), 116.2 (Cheb-C-6″), 110.7 (Gal-C-2″, 6″), 110.7 (Gal-C-2′, 6′), 110.2 (Gal-C-2‴, 6‴), 92.9 (C-1), 76.2 (C-5), 71.9 (C-2), 69.5 (C-4), 67.1 (Cheb-C-2′), 65.2 (C-6), 62.9 (C-3), 41.8 (Cheb-C-3′), 40.2 (Cheb-C-4′), 30.8 (Cheb-C-5′)

1,2,3,4,6-Penta-*O*-galloyl-*β*-D-glucose (**8**)

White amorphous powder

$C_{41}H_{32}O_{26}$

$[\alpha]_D^{20}$= 13.3 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.9), 281 (4.7) nm

IR ν_{max}: 3373, 1699, 1513, 1457, 1341, 1209 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 939.1119 $[M - H]^-$ (calcd for $C_{41}H_{31}O_{26}$, 939.1104)

^{1}H (400 MHz, DMSO-d_6) NMR: δ 6.97 (2H, *s*, Gal-H-2‴, 6‴), 6.91 (2H, *s*, Gal-H-2‴, 6‴), 6.84 (2H, *s*, Gal-H-2″, 6″), 6.81 (2H, *s*, Gal-H-2⁗, 6⁗), 6.76 (2H, *s*, Gal-H-2′, 6′), 6.37 (1H, *d*, *J* = 8.2 Hz, H-1), 5.95 (1H, *t*, *J* = 9.6 Hz, H-3), 5.43 (2H, *m*, H-2, 4), 4.59 (1H, *d*, *J* = 9.8 Hz, H-5), 4.30 (2H, br *s*, H-6)

^{13}C (100 MHz, DMSO-d_6) NMR: δ 165.40 (Gal-C-7⁗), 164.79 (Gal-C-7‴), 164.57 (Gal-C-7⁗), 164.44 (Gal-C-7″), 163.91 (Gal-C-7′), 145.66 (Gal-C-3‴, 5⁗), 145.54 (Gal-C-3‴, 5‴), 145.48 (Gal-C-3″, 5″), 145.44 (Gal-C-3⁗, 5⁗), 145.36 (Gal-C-3′, 5′), 139.61 (Gal-C-4⁗), 139.11 (Gal-C-4‴), 139.09 (Gal-C-4″), 138.88 (Gal-C-4⁗), 138.74 (Gal-C-4′), 118.87 (Gal-C-1⁗), 118.08 (Gal-C-1‴), 118.05 (Gal-C-1″), 117.91 (Gal-C-1⁗), 117.34 (Gal-C-1′), 109.00 (Gal-C-2⁗,

6‴′), 108.87 (Gal-C-2‴, 6‴), 108.77 (Gal-C-2″, 6″), 108.72 (2C) (Gal-C-2′, 6′, Gal-C-2‴′, 6‴′), 91.66 (C-1), 72.11 (C-5), 71.90 (C-3), 70.55 (C-2), 67.73 (C-4), 61.44 (C-6)

4-*O*-(4″-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**9**)

Brown amorphous powder

$C_{27}H_{20}O_{16}$

$[\alpha]_D^{20}$= -149.0 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (4.7) 258 (4.8), 356 (4.1) nm

IR ν_{max}: 3396, 1711, 1613, 1500, 1451, 1341, 1231 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 599.0684 $[M - H]^-$ (calcd for $C_{27}H_{19}O_{16}$, 599.0673)

^1H (400 MHz, CD_3OD) NMR: See Table 3

^{13}C (100 MHz, CD_3OD) NMR: See Table 3

4-*O*-(3″,4″-Di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**10**)

Brown amorphous powder

$C_{34}H_{24}O_{20}$

$[\alpha]_D^{20}$= -60.5 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.7) 256 (4.7), 358 (4.0) nm

IR ν_{max}: 3389, 1702, 1615, 1453, 1342, 1219 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 751.0787 $[M - H]^-$ (calcd for $C_{34}H_{23}O_{20}$, 751.0783)

^1H (400 MHz, CD_3OD) NMR: See Table 3

^{13}C (100 MHz, CD_3OD) NMR: See Table 3

4-*O*-(2″,4″-Di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**11**)

Brown amorphous powder

$C_{34}H_{24}O_{20}$

$[\alpha]_D^{20}$= -37.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (4.6) 258 (4.5), 356 (3.9) nm

IR ν_{max}: 3364, 1708, 1615, 1448, 1347, 1213, 1028 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 751.0786 $[M - H]^-$ (calcd for $C_{34}H_{23}O_{20}$, 751.0783)

^1H (600 MHz, CD_3OD) NMR: See Table 3

^{13}C (150 MHz, CD_3OD) NMR: See Table 3

1,2,3,6-Tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**)

Brown amorphous powder

$C_{43}H_{34}O_{23}$

$[\alpha]_D^{20}$= 28.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 281 (4.7) nm

IR ν_{max}: 3411, 1708, 1617, 1532, 1455, 1321, 1208, 1031 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 917.1425 $[M - H]^-$ (calcd for $C_{43}H_{33}O_{23}$, 917.1413)

^1H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 4

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 4

1,2,3-Tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (**13**)

Brown amorphous powder

$C_{36}H_{30}O_{19}$

$[\alpha]_D^{20}$ = 52.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.9), 280 (4.7) nm

IR ν_{max}: 3412, 1707, 1617, 1532, 1455, 1321, 1205, 1031 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 765.1323 $[M - H]^-$ (calcd for $C_{36}H_{29}O_{19}$, 765.1303)

^1H (800 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 4

^{13}C (200 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 4

Table 3. ^{1}H and ^{13}C NMR data for compounds **9-11**

Position	**9**[a]		**10**[a]		**11**[b]	
	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C
α-L-Rhamnose						
1″	5.65, *s*	101.6	5.70, *s*	101.7	5.76, *s*	99.0
2″	4.29, br *s*	71.9	4.58, br *s*	69.7	5.65, br *s*	73.7
3″	4.30, *m*	70.3	5.66, *m*	73.6	4.61, *d* (8.3)	68.9
4″	5.25, *t* (9.5)	75.3	5.60, *m*	72.2	5.37, *t* (9.8)	75.4
5″	4.04, *dq* (12.4, 6.1)	69.4	4.25, *m*	69.4	4.19, *m*	69.5
6″	1.19, *d* (6.2)	18.0	1.29, *d* (6.1)	18.0	1.26, *d* (6.1)	18.1
Ellagic acid						
1		116.3		116.2		116.6
2		137.9		137.7		138.0
3		142.9		143.0		143.5
4		147.7		147.6		147.6
5	7.89, *s*	113.6	7.87, *s*	113.6	7.92, *s*	113.9
6		109.2		108.9		108.9
7		161.1		161.0		161.2
1′		113.3		113.1		113.3
2′		138.0		137.7		138.0
3′		141.0		141.1		141.2
4′		150.1		150.1		150.2
5′	7.51, *s*	111.9	7.44, *s*	111.9	7.52, *s*	111.8
6′		110.0		109.7		109.9
7′		161.0		160.9		161.1
	4″-*O*-Galloyl		3″-*O*-Galloyl		2″-*O*-Galloyl	
1‴		121.4		120.8		120.9
2‴, 6‴	7.10, *s*	110.3	7.08, *s*	110.5	7.21, *s*	110.5
3‴, 5‴		146.5		146.4		146.5
4‴		139.9		140.2		140.2
7‴		168.1		168.2		167.7
			4″-*O*-Galloyl		4″-*O*-Galloyl	
1⁗				120.8		121.2
2⁗, 6⁗			7.02, *s*	110.2	7.12, *s*	110.3
3⁗, 5⁗				146.4		146.5
4⁗				140.1		140.0
7⁗				167.7		168.0

^{1}H NMR data were measured at [a]400 and [b]600 MHz in CD_3OD, respectively. ^{13}C NMR data were measured at [a]100 and [b]150 MHz in CD_3OD, respectively.

Table 4. ^{1}H and ^{13}C NMR data for compounds **12-13**

Position	**12**[a]		**13**[b]	
	δ_H	δ_C	δ_H	δ_C
β-D-Glucose				
1	6.25, *d* (8.3)	93.3	6.18, *d* (8.2)	93.4
2	5.60, *dd* (8.3, 9.7)	71.7	5.47, *dd* (8.2, 9.6)	71.8
3	5.94, *t* (9.7)	73.3	5.66, *t* (9.4)	75.9
4	5.61, *t* (9.6)	69.6	4.07, *t* (9.4)	69.3
5	4.53, *ddd* (2.5, 4.6, 10.0)	73.5	4.14, *ddd* (1.7, 5.0, 9.5)	75.9
6a	4.57, *dd* (2.5, 12.4)	63.0	4.60, *dd* (1.7, 12.0)	63.8
6b	4.37, *dd* (4.6, 12.4)		4.50, *dd* (5.0, 12.0)	
	1-*O*-Galloyl		1-*O*-Galloyl	
1′		119.3		120.2
2′, 6′	7.06, *s*	110.2	7.11, *s*	110.3
3′, 5′		146.1		146.0
4′		139.6		139.6
7′		165.3		165.0
	2-*O*-Galloyl		2-*O*-Galloyl	
1″		119.8[d]		120.8
2″, 6″	6.99[c], *s*	110.0[e]	7.01, *s*	110.1
3″, 5″		145.9[f]		145.9
4″		139.5[g]		139.1
7″		166.3[h]		165.8
	3-*O*-Galloyl		3-*O*-Galloyl	
1‴		120.0[d]		121.4
2‴, 6‴	6.99[c], *s*	110.0[e]	7.08, *s*	110.1
3‴, 5‴		146.0[f]		145.8
4‴		139.4[g]		138.8
7‴		166.4[h]		166.2
	6-*O*-Galloyl			
1⁗		120.8		
2⁗, 6⁗	7.13, *s*	109.9		
3⁗, 5⁗		146.0		
4⁗		139.1		
7⁗		166.7		
	4-*O*-Cinnamoyl		6-*O*-Cinnamoyl	
1′		134.8		135.4
2′, 6′	7.57, *m*	129.1	7.71, *m*	129.2
3′, 5′	7.36, *m*	129.7	7.45, *m*	129.8
4′	7.37, *m*	131.5	7.43, *m*	131.3
7′	7.63, *d* (16.0)	147.7	7.72, *d* (16.0)	145.8
8′	6.45, *d* (16.0)	117.3	6.62, *d* (16.0)	118.7
9′		166.2		167.1

^{1}H NMR data were measured at ^{a}600 and ^{b}800 MHz in acetone-d_6:D_2O = 9:1, respectively. ^{13}C NMR data were measured at ^{a}150 and ^{b}200 MHz in acetone-d_6:D_2O = 9:1, respectively. $^{c–h}$ Exchangable

2.4. Further isolation of constituents of *T. chebula*

The B1 fraction (40 g) was subjected to was subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (1:99 to 25:75) to yield five fractions (B1a-B1f). Fraction B1a was separated on semi-preparative HPLC column (MeCN/H_2O with 0.1% formic acid 1:99) to give compound **37** (10 mg) and **34** (150 mg). Fraction B1b was recrystallized using water to afford **27** (2.0 g). Fraction B1c subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (1:99 to 15:85) to yield nine fractions (B1c1-B1c9). Fraction B1c3 was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 3:97) to give compound **31** (5 mg). Compound **29** (51 mg), **30** (32 mg), **44** (15 mg), and **46** (90 mg) were isolated from B1c5 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 4:96). Compound **19** (110 mg) and **41** (29 mg) were obtained from B1c6 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 5:95). B1c7 was separated on semi-preparative HPLC column (MeCN/H_2O with 0.1% formic acid 5:95) to give compound **14** (20 mg), **32** (22 mg), and **47** (6 mg). Compound **48** (120 mg) and **35** (33 mg) were obtained from B1c8 and B1c9 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 6:94 and 5:95), respectively. Fraction B1d subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (3:97 to 20:80) to yield seven fractions (B1d1-B1d7). Compound **49** (55 mg) and **15** (30 mg) were isolated from B1d2 and B1d3 by recrystallization (Water), respectively. B1d5 was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield five subfractions (B1d5a-

B1d5e). Compound **42** (700 mg) was obtained from B1d5d. B1d5a was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 8:92) to give compound **25** (70 mg). Compound **33** (150 mg) and **45** (10 mg) were obtained from B1d5b and B1d5e through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 7:93), respectively. Compound **38** (4 mg) and **40** (10 mg) were purified from B1d7 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 7:93). Fraction B1e was subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (5:80 to 20:80) to yield eight fractions (B1e1-B1e8). B1e3 was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 6:94) to give compound **28** (50 mg). Compound **36** (8 mg) was isolated from B1e4 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 8:92). Compound **20** (44 mg) and **22** (29 mg) were obtained from B1e5 and B1e8 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 9:91 and 10:90), respectively.

The B2 was subjected to MPLC eluted with a MeCN/H_2O with 0.1% formic acid gradient (5:95 to 30:70) to yield six subfractions (B2a-B2f). Fraction B2d was separated by semi-preparative HPLC, eluted with a MeCN/ H_2O with 0.1% formic acid (8:92 to 15:85) to yield 2 subfractions (B2d1 and B2d2). Compound **43** (18 mg) was obtained from B2d1 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 9:91). Fraction B2e was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield eight subfractions (B2e1-B2e8). B2e6 was recrystallized using water to afford **26** (6 mg). Compound **18** (50 mg) and **21** (15 mg) were isolated from B2e8 through semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 10:90 to 15:85). Compound **23** (15 mg) and **24** (31 mg) were obtained from B2g by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 15:85).

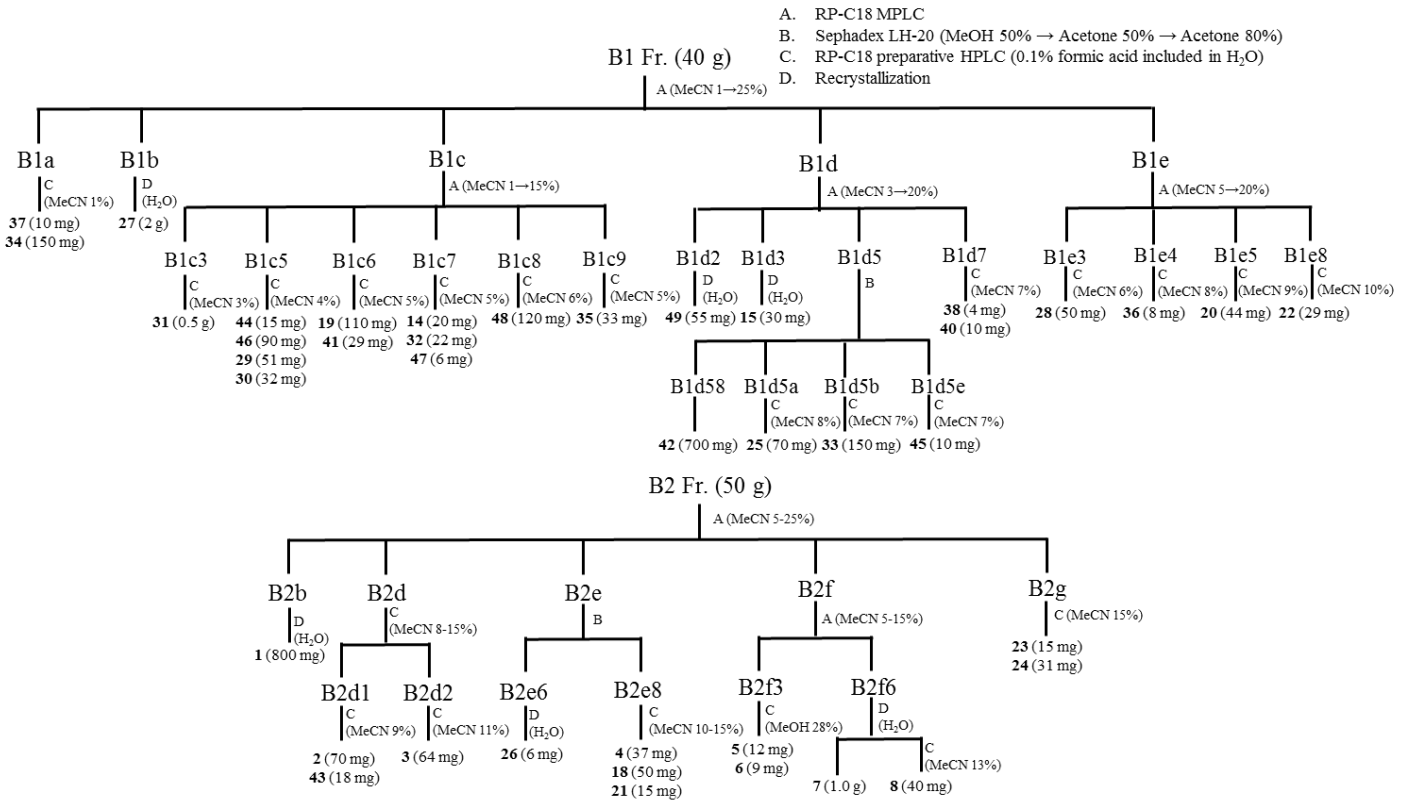
B3 fraction was subjected to MPLC eluted with a MeOH/H_2O with 0.1% formic

acid gradient (20:80 to 50:50) to yield seven subfractions (B3a-B3g). B3a was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) and purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 25:75) to give compound **52** (2 mg), **57** (4 mg), and **66** (2 mg). B3b was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield eight subfractions (B3d1-B3d8). Compound **37** (1.2 g) was isolated from B3d4 by recrystallization (MeOH). Compound **16** (14 mg) and **17** (11 mg) were obtained from B3d8 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 26:74). Fraction B3f was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield five subfractions (B3f1-B3f5). Compound **54** (400 mg) was isolated from B3f1 by recrystallization (MeOH). Compound **55** (27 mg) was obtained from B3f2 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 24:76). Fraction B3g was separated on Sephadex LH-20 column (MeOH-H_2O 5:5 to acetone-H_2O 8:2) to yield five subfractions (B3g1-B3g5). Compound **59** (20 mg), **63** (9 mg), and **65** (3 mg) were isolated from B3g1 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 29:71). B3g2 was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 30:70) to give compound **61** (8 mg).

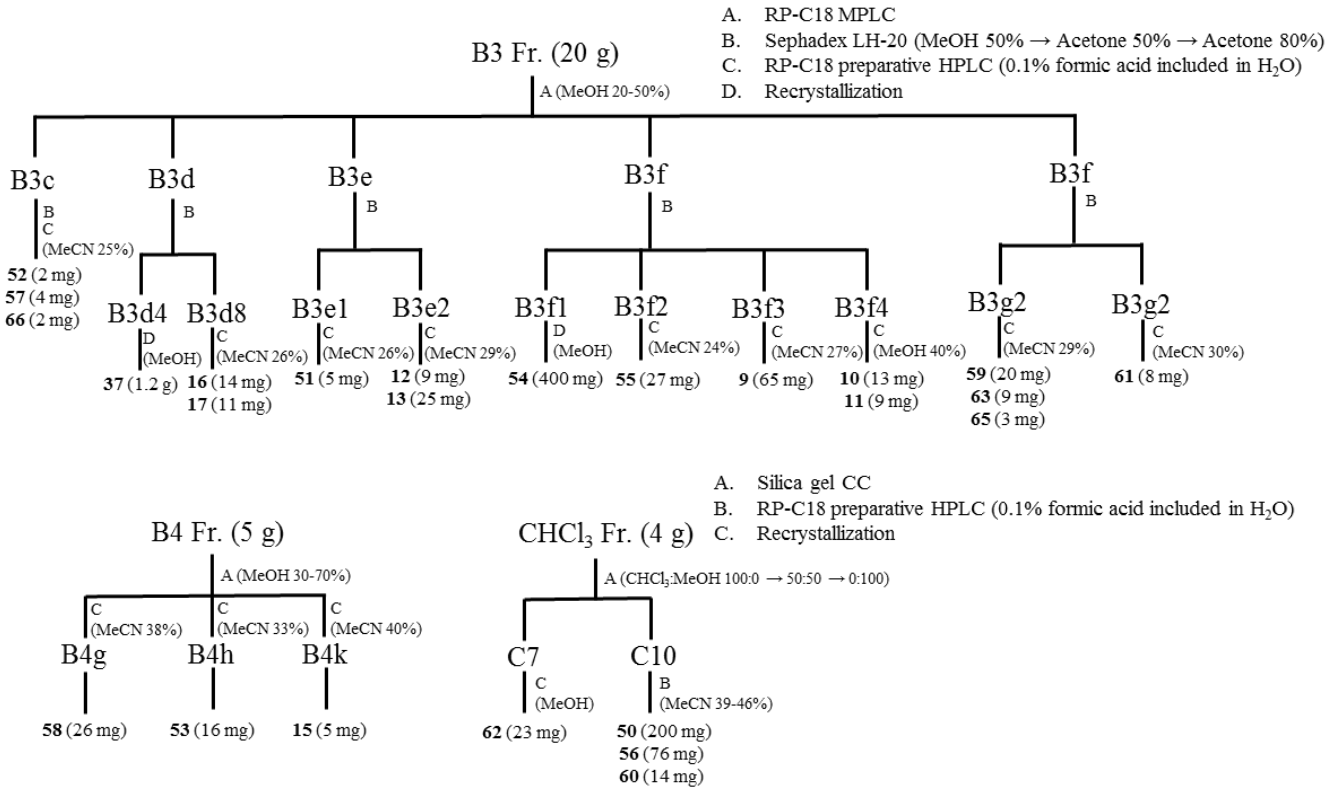
Fraction B4 was subjected to MPLC eluted with a MeOH/H_2O with 0.1% formic acid gradient (30:70 to 70:30) to yield eleven subfractions (B4a-B4k). B4g was purified by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 38:62) to give compound **58** (26 mg). Compound **53** (16 mg) and **64** (5 mg) were obtained from B4h and B4k by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 33:67 and 40:60), respectively.

The $CHCl_3$ fraction was subjected to silica gel column chromatography eluted with mixtures of $CHCl_3$:MeOH (100:0 → 50:50 → 0:100) to yield eleven fractions (C1-C11). Compound **62** (23 mg) was isolated from by recrystallization (MeOH).

Compound **50** (200 mg), **56** (76 mg), and **60** (14 mg) were obtained from C10 by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 39:61 to 46:54).



Scheme 5. Isolation of compounds from *T. chebula* fruits (Continued)



Scheme 6. Isolation of compounds from *T. chebula* fruits

1,3-Di-*O*-galloyl-*β*-D-glucose (**14**)

White amorphous powder

$C_{20}H_{20}O_{14}$

$[\alpha]_D^{20}$= 4.7 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.9), 278 (4.5) nm

IR ν_{max}: 3390, 1701, 1617, 1537, 1455, 1341, 1213 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 483.0777 $[M - H]^-$ (calcd for $C_{20}H_{19}O_{14}$, 483.0775)

^1H (300 MHz, CD_3OD) NMR: *δ* 7.14, 7.12 (each 2H, *s*, Gal-H-2′, 6′, Gal-H-2″, 6″), 5.79 (1H, *d*, *J* = 8.2 Hz, H-1), 5.24 (1H, *t*, *J* = 9.4 Hz, H-3), 3.87 (1H, *dd*, *J* = 2.1, 12.2 Hz, H-6a), 3.76 (1H, *m*, H-4) 3.73 (1H, *m*, H-6b), 3.70 (1H, *m*, H-2), 3.56 (1H, *ddd*, J = 2.2, 4.6, 9.9 Hz, H-5)

^{13}C (75 MHz, CD_3OD) NMR: *δ* 168.1, 166.8 (Gal-C-7′, Gal-C-7″), 146.5, 146.3 (Gal-C-3′, 5′, Gal-C-3″, 5″), 140.3, 139.7 (Gal-C-4′, Gal-C-4″), 121.6, 120.5 (Gal-C-1′, Gal-C-1″), 110.5, 110.3 (Gal-C-2′, 6′, Gal-C-2″, 6″), 95.8 (C-1), 79.1 (C-3), 78.7 (C-5), 72.6 (C-2), 69.3 (C-4), 61.9 (C-6)

1,6-Di-*O*-galloyl-*β*-D-glucose (**15**)

White amorphous powder

$C_{20}H_{20}O_{14}$

$[\alpha]_D^{20}$= -29.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.8), 278 (4.5) nm

IR ν_{max}: 3390, 1701, 1618, 1513, 1456, 1342, 1220 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 483.0782 $[M - H]^-$ (calcd for $C_{20}H_{19}O_{14}$ 483.0775)

^1H (300 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 7.16, 7.12 (each 2H, *s*, Gal-H-2′, 6′, Gal-H-2″, 6″), 5.73 (1H, *d*, *J* = 7.7 Hz, H-1), 4.55 (1H, *dd*, *J* = 1.9, 12.2 Hz, H-6a), 4.38 (1H, *dd*, *J* = 5.0, 12.2 Hz, H-6b), 3.79 (1H, m, H-5), 3.56–3.49 (3H, *m*, H-2, H-3, H-4)

^{13}C (75 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 166.7, 165.5 (Gal-C-7′, Gal-C-7″), 146.0, 146.0 (Gal-C-3′, 5′, Gal-C-3″, 5″), 139.3, 138.8 (Gal-C-4′, Gal-C-4″), 121.4, 120.70 (Gal-C-1′, Gal-C-1″), 110.20, 109.80 (Gal-C-2′, 6′, Gal-C-2″, 6″), 95.6 (C-1), 77.7 (C-3), 75.8 (C-5), 73.8 (C-2), 70.80 (C-4), 64.14 (C-6)

1,6-Di-*O*-galloyl-2-*O*-cinnamoyl-*β*-D-glucose (**16**)

Light brown amorphous powder

$C_{29}H_{26}O_{15}$

$[\alpha]_D^{20}$= -56.0 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.8), 280 (4.5) nm

IR ν_{max}: 3382, 1701, 1513, 1456, 1341, 1212 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 613.1219 $[M - H]^-$ (calcd for $C_{29}H_{25}O_{15}$ 613.1193)

^1H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: 7.72 (1H, *d*, *J* = 16.0 Hz, Cin-H-7′), 7.61 (2H, m, Cin-H-2′, 6′), 7.38 (3H, m, Cin-H-3′, Cin-H-4′, Cin-H-5′) 7.11, 7.05 (each 2H, Gal-H-2′, 6′, Gal-H-2″, 6″), 6.53 (1H, *d*, *J* = 16.0 Hz, Cin-H-8″) 5.86 (1H, *d*, *J* = 8.3 Hz, H-1), 5.18 (1H, *t*, *J* = 8.0 Hz, H-2), 4.60 (1H, *dd*, *J* = 1.6, 12.1 Hz, H-6a), 4.37 (1H, *dd*, *J* = 5.4, 12.1 Hz, H-6b), 3.92 (3H, m, H-3, H-4, H-5)

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: 167.1, 167.0 (Gal-C-7′, Gal-C-7″), 165.6 (Cin-C-9′), 146.4 (Cin-C-7′) 146.1, 145.9 (Gal-C-3′, 5′, Gal-C-3″, 5″), 139.7, 139.0 (Gal-C-4′, Gal-C-4″), 134.9 (Cin-C-1′), 131.4 (Cin-C-4′), 129.72 (Cin-C-3′, 5′), 129.03 (Cin-C-2′, 6′), 121.0, 119.5 (Gal-C-1′, Gal-C-1″), 118.1 (Cin-C-8′), 110.0, 109.7 (Gal-C-2′, 6′, Gal-C-2″, 6″), 93.5 (C-1), 75.9 (C-5), 74.8 (C-3), 73.7 (C-2), 70.9 (C-4), 64.0 (C-6)

1,2-Di-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (**17**)

Brown amorphous powder

$C_{29}H_{26}O_{15}$

$[\alpha]_D^{20}$= -66.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.8), 279 (4.6) nm

IR ν_{max}: 3424, 1701, 1618, 1513, 1455, 1338, 1207 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 613.1210 $[M - H]^-$ (calcd for $C_{29}H_{25}O_{15}$. 613.1193)

^{1}H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 7.67 (1H, *d*, *J* = 16.0 Hz, Cin-H-7′) 7.65 (2H, m, Cin-H-2′, 6′), 7.39 (3H, m, Cin-H-3′, Cin-H-4′, Cin-H-5′), 7.06 (2H, *s*, Gal-H-2″, 6″), 7.03 (2H, *s*, Gal-H-2′, 6′), 6.55 (1H, *d*, *J* = 16.0 Hz, Cin-H-8″), 5.89 (1H, *d*, *J* = 8.3 Hz, H-1), 5.19 (1H, *dd*, *J* = 8.3, 9.5 Hz, H-2), 4.54 (1H, *dd*, *J* = 2.0, 12.1 Hz, H-6a), 4.34 (1H, *dd*, *J* = 5.7, 12.1 Hz, H-6b) 3.93 (1H, *t*, *J* = 9.2 Hz, H-3), 3.87 (1H, *ddd*, *J* = 2.1, 5.7, 9.9, H-5) 3.64 (1H, *t*, *J* = 8.0 Hz, H-4)

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 166.7 (Cin-C-9′), 166.1 (Gal-C-7″), 164.7 (Gal-C-7′), 145.4 (Gal-C-3′, 5′), 145.2 (Gal-C-3″, 5″), 145.2 (Cin-C-7′), 139.0 (Gal-C-4′), 138.5 (Gal-C-4″), 134.4 (Cin-C-1′), 130.6 (Cin-C-4′), 129.1 (Cin-

C-3′, 5′), 128.4 (Cin-C-2′, 6′), 120.0 (Gal-C-1″), 119.0 (Gal-C-1′), 117.7 (Cin-C-8′), 109.3 (Gal-C-2′, 6′), 109.2 (Gal-C-2″, 6″), 92.7 (C-1), 75.1 (C-5), 74.3 (C-3), 73.1 (C-2), 70.2 (C-4), 63.4 (C-6).

Methyl chebulagate (**18**)

Light brown amorphous powder

$C_{42}H_{32}O_{27}$

$[\alpha]_D^{20}$= -7.7 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 222 (4.9), 279 (4.5) nm

CD (MeOH): 239 (Δε –16.57), 262 (Δε 5.13), 286 (Δε -14.08)

IR ν_{max}: 3437, 1725, 1516, 1324, 1210, 1033 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 967.1055 $[M - H]^-$ (calcd for $C_{42}H_{31}O_{27}$ 967.1053)

^1H (600 MHz, acetone-d_6) NMR:, 7.51 (1H, *s*, Cheb-H-2″), 7.19 (1H, *s*, Gal-H-2′, 6′), 6.08 (1H, *s*, HHDP-H-5′), 6.67 (1H, *s*, HHDP-H-5″), 6.51 (1H, *s*, H-1), 5.90 (1H, *s*, H-3), 5.51 (1H, *s*, H-2), 5.23 (1H, *d*, *J* = 3.5 Hz, H-4), 5.13 (1H, *dd*, *J* = 0.8, 7.3 Hz, Cheb-H-3′), 4.97 (1H, *d*, Cheb-H-2′), 4.82 (1H, *m*, H-5), 4.78 (1H, *m*, H-6a), 4.41 (1H, *dd*, *J* = 7.4, 10.2 Hz, H-6b), 3.90 (1H, *t*, *J* = 8.2 Hz, Cheb-H-4′), 3.54 (3H, *s*, OMe) 2.17 (2H, br *d*, Cheb-H-5′)

^{13}C (150 MHz, acetone-d_6) NMR: δ 173.5 (Cheb-C-7′), 172.2 (Cheb-C-6′), 169.6 (Cheb-C-1′), 168.3 (HHDP-C-7″), 166.2 (HHDP-C-7′), 165.4 (Cheb-C-7″), 164.9 (Gal-H-7′), 146.5 (Cheb-C-3″), 145.9 (Gal-C-3′, 5′), 145.4 (HHDP-C-4″), 145.2 (HHDP-C-2′), 145.0 (HHDP-C-2″), 144.1 (HHDP-C-4′), 141.0 (Cheb-C-5″), 139.7 (Gal-C-4′), 139.4 (Cheb-C-4″), 137.7 (HHDP-C-3′), 136.5 (HHDP-C-3″), 125.7

(HHDP-C-6″), 124.8 (HHDP-C-6′), 120.3 (Gal-C-1′), 119.2 (Cheb-C-1″), 117.2 (Cheb-C-2″), 117.0 (HHDP-C-1′) 115.1 (HHDP-C-1″), 115.9 (Cheb-C-6″), 110.8 (Gal-C-2′, 6′), 110.3 (HHDP-C-5′), 108.0 (HHDP-C-5″), 91.6 (C-1), 73.7 (C-5), 70.6 (C-2), 66.6 (Cheb-C-2′), 66.4 (C-4) , 64.0 (C-6), 61.9 (C-3), 52.0 (OMe) 41.1 (Cheb-C-3′), 39.3 (Cheb-C-4′), 30.2 (Cheb-C-5′)

Phyllanemblinin E (**19**)

Yellowish amorphous powder

$C_{27}H_{26}O_{20}$

$[\alpha]_D^{20}$= 1.6 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 221 (4.7), 278 (4.3) nm

IR ν_{max}: 3423, 1721, 1617, 1513, 1455, 1339, 1206 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 669.0928 $[M - H]^-$ (calcd for $C_{27}H_{25}O_{20}$, 669.0939)

^1H (500 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 5

^{13}C (125 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 6

1′-*O*-Methyl neochebulanin (**20**)

White amorphous powder

$C_{28}H_{28}O_{20}$

$[\alpha]_D^{20}$= 15.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.7), 284 (4.3) nm

CD (MeOH): 214 (Δε -4.77), 235 (Δε 5.07), 261 (Δε -0.26), 291 (Δε 1.93)

IR ν_{max}: 3420, 2919, 1732, 1540, 1339, 1199, 1060, 669 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 683.1098 $[M - H]^-$ (calcd for $C_{28}H_{27}O_{20}$ 683.1098)

^1H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 5

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 6

Dimethyl neochebulinate (**21**)

Yellowish amorphous powder

$C_{43}H_{38}O_{28}$

$[\alpha]_D^{20}$= 28.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (4.9), 281 (4.5) nm

CD (MeOH): 208 (Δε -16.80), 227 (Δε 5.59), 235 (sh, Δε 4.26), 267 (Δε -2.23), 293 (Δε 5.23)

IR ν_{max}: 3411, 2923, 1723, 1530, 1322, 1209, 1038, 660 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1001.1484 $[M - H]^-$ (calcd for $C_{43}H_{37}O_{28}$ 1001.1472)

^1H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 5

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 6

Neochebulagic acid (**22**)

Yellowish amorphous powder

$C_{41}H_{32}O_{28}$

$[\alpha]_D^{20}$= -49.0 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (5.0), 281 (4.6) nm

IR ν_{max}: 3373, 1710, 1513, 1457, 1342 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 971.0989 $[M - H]^-$ (calcd for $C_{41}H_{31}O_{28}$ 971.1002)

^1H (500 MHz, CD_3OD) NMR: See Table 7

^{13}C (125 MHz, CD_3OD) NMR: See Table 8

Dimethyl neochebulagate (**23**)

Yellowish amorphous powder

$C_{43}H_{36}O_{28}$

$[\alpha]_D^{20}$= -1.7 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 222 (5.0), 280 (4.6) nm

CD (MeOH): 240 (Δε -27.35), 263 (Δε 10.41), 285 (Δε -18.37)

IR ν_{max}: 3420, 2949, 1717, 1526, 1361, 1198, 1032, 669 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 999.1298 $[M - H]^-$ (calcd for $C_{43}H_{35}O_{28}$ 999.1315)

^1H (600 MHz, acetone-d_6) NMR: See Table 7

^{13}C (150 MHz, acetone-d_6) NMR: See Table 8

Dimethyl 4′-*epi*-neochebulagate (**24**)

Yellowish amorphous powder

$C_{43}H_{36}O_{28}$

$[\alpha]_D^{20}$= 21.9 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 221 (4.9), 280 (4.5) nm

CD (MeOH): 241 (Δε -19.94), 263 (Δε 7.83), 285 (Δε -12.08)

IR ν_{max}: 3437, 2964, 1725, 1515, 1329, 1204, 1040, 677 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 999.1315 $[M - H]^-$ (calcd for $C_{43}H_{35}O_{28}$ 999.1315)

^1H (600 MHz, acetone-d_6) NMR: See Table 7

^{13}C (150 MHz, acetone-d_6) NMR: See Table 8

Chebulanin (**25**)

White amorphous powder

$C_{27}H_{24}O_{19}$

$[\alpha]_D^{20}$= 5.3 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 222 (4.6), 280 (4.3) nm

IR ν_{max}: 3385, 1701, 1513, 1457, 1341 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 651.0825 $[M - H]^-$ (calcd for $C_{27}H_{23}O_{19}$, 651.0834)

^1H (300 MHz, CD_3OD) NMR: δ 7.45 (1H, *s*, Cheb-H-2″), 7.13 (1H, *s*, Gal-H-2′, 6′), 6.37 (1H, *d*, *J* = 2.8 Hz, H-1), 5.22 (1H, *m*, H-2), 5.10 (1H, *dd*, *J* = 1.5, 7.2 Hz, Cheb-H-3′), 4.82 (1H, *m*, H-4), 4.81 (1H, *m*, H-3), 4.78 (1H, *m*, Cheb-H-2′) 4.30 (1H, *dt*, *J* = 12.4, 6.4 Hz, H-5), 4.04 (2H, m, H-6), 3.81 (1H, *ddd*, *J* = 1.0, 4.1, 5.5

Hz, Cheb-H-4′), 2.15 (2H, *m*, Cheb-H-5′)

^{13}C (75 MHz, CD_3OD) NMR: δ 174.8 (Cheb-C-6′), 174.6 (Cheb-C-7′), 170.7 (Cheb-C-1′), 166.6 (Gal-C-7′), 166.4 (Cheb-C-7″), 147.1 (Cheb-C-3″), 146.5 (Gal-C-3′, 5′), 141.2 (Cheb-C-5″), 140.4 (Gal-C-4′), 140.0 (Cheb-C-4″), 120.4 (Gal-C-1′), 119.4 (Cheb-C-1″), 117.3 (Cheb-C-2″), 115.9 (Cheb-C-6″), 110.4 (Gal-C-2′, 6′), 92.9 (C-1), 79.4 (C-5), 74.1 (C-2), 72.0 (C-4), 67.0 (Cheb-C-2′), 63.5 (C-6), 61.6 (C-3), 41.4 (Cheb-C-3′), 39.9 (Cheb-C-4′), 30.1 (Cheb-C-5′)

Eschweilenol C (**26**)

Brown amorphous powder

$C_{20}H_{16}O_{12}$

$[\alpha]_D^{20}$= -68.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 255 (4.6), 360 (4.0) nm

IR ν_{max}: 3331, 1738, 1617, 1501, 1541, 1341 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 447.0561 $[M - H]^-$ (calcd for $C_{20}H_{15}O_{12}$ 447.0564)

^{1}H (400 MHz, DMSO-d_6) NMR: δ 7.70 (1H, *s*, H-5), 7.43 (1H, *s*, H-5′), 5.42 (1H, *s*, H-1″), 3.96 (1H, *s*, H-2″), 3.80 (1H, *dd*, *J* = 3.0, 9.3 Hz, H-3″), 3.50 (1H, *m*, H-5″), 3.29 (1H, *t*, *J* = 9.4 Hz, H-4″), 1.10 (3H, *d*, *J* = 6.2 Hz, H-6″)

^{13}C (100 MHz, DMSO-d_6) NMR: δ 159.2 (C-7), 159.0 (C-7′), 148.8 (C-4′), 146.4 (C-4), 141.4 (C-3), 140.0 (C-3′), 136.7 (C-2), 136.4 (C-2′), 114.6 (C-6), 111.9 (C-6′), 111.6 (C-5), 110.2 (C-5′), 107.9 (C-1), 107.3 (C-1′), 100.2 (C-1″), 71.8 (C-4″), 70.1 (C-5″), 70.0 (C-3″), 69.9 (C-2″), 17.9 (C-6″)

Gallic acid (**27**)

White amorphous powder

$C_7H_6O_5$

UV (MeOH) λ_{max} (log ε): 218 (4.4), 273 (4.0) nm

IR ν_{max}: 3373, 1693, 1513, 1457, 1335 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 169.0131 $[M - H]^-$ (calcd for $C_7H_5O_5$ 169.0137)

^1H (300 MHz, DMSO-d_6) NMR: δ 6.91 (2H, *s*, H-2, 6)

^{13}C (75 MHz, DMSO-d_6) NMR: δ 167.5 (C-7), 145.4 (C-3, 5), 138.0 (C-4), 120.4 (C-1), 108.7 (C-2, 6)

Methyl gallate (**28**)

Yellowish amorphous powder

$C_8H_8O_5$

UV (MeOH) λ_{max} (log ε): 219 (4.4), 275 (4.1) nm

IR ν_{max}: 3373, 1698, 1513, 1458, 1252 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 183.0289 $[M - H]^-$ (calcd for $C_8H_7O_5$ 183.0293)

^1H (300 MHz, CD_3OD) NMR: δ 7.04 (2H, *s*, H-2, 6), 3.81 (3H, *s*, OMe)

^{13}C (75 MHz, CD_3OD) NMR: δ 168.9 (C-7), 146.5 (C-3, 5), 139.7 (C-4), 121.4 (C-1), 110.0 (C-2, 6), 52.3 (OMe)

4-*O*-Galloyl-(-)-shikimic acid (**29**)

Yellowish syrup

$C_{14}H_{14}O_9$

$[\alpha]_D^{20}$= -224.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.7), 278 (4.3) nm

IR ν_{max}: 3468, 1698, 1618, 1514, 1455, 1316, 1244 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 325.0560 $[M - H]^-$ (calcd for $C_{14}H_{13}O_9$ 325.0560)

^1H (300 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 9

^{13}C (75 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 9

5-*O*-Galloyl-(-)-shikimic acid (**30**)

Yellowish syrup

$C_{14}H_{14}O_9$

$[\alpha]_D^{20}$= -165.99 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.6), 278 (4.2) nm

IR ν_{max}: 3473, 1700, 1513, 1457, 1342 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 325.0559 $[M - H]^-$ (calcd for $C_{14}H_{13}O_9$ 325.0560)

^1H (300 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 9

^{13}C (75 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 9

6-*O*-Galloyl-D-glucose (**31**)

Yellowish amorphous powder

$C_{13}H_{16}O_{10}$

$[\alpha]_D^{20}$= 13.6 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.3), 275 (3.9) nm

IR ν_{max}: 3390, 1700, 1514, 1456, 1342, 1235 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 331.0649 $[M - H]^-$ (calcd for $C_{13}H_{15}O_{10}$ 331.0665)

^1H (300 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 10

^{13}C (75 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 167.5 (*α, β*-C-7′), 145.4 (*α, β*-C-3′, 5′), 138.0 (*α, β*-C-4′), 120.4 (*α, β*-C-1′), 108.7 (*α, β*-C-2′, 6′) 97.7 (*β*-C-1), 93.2 (*α*-C-1), 77.2 (*β*-C-3), 75.6 (*β*-C-5), 74.6 (*α*-C-3), 74.2 (*β*-C-2), 73.1 (*α*-C-5), 71.1 (*α*-C-2), 70.9 (*β*-C-4), 70.2(*α*-C-4), 64.5 (*α, β*-C-6)

3,6-Di-*O*-galloyl-D-glucose (**32**)

Yellowish amorphous powder

$C_{20}H_{20}O_{14}$

$[\alpha]_D^{20}$= 61.9 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.7), 276 (4.3) nm

IR ν_{max}: 3373, 1700, 1650, 1514, 1457, 1342,1230 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 483.0777 $[M - H]^-$ (calcd for $C_{20}H_{19}O_{14}$ 483.0775)

^1H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 10

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 167.6, 167.2 (*α, β*-C-7′, 7″), 145.9, 145.8 (*α, β*-C-3′,5′, 3″,5″), 138.9, 138.7 (*α, β*-C-4′, 4″), 121.7, 121.3 (*α, β*-C-1′, 1″), 109.9, 109.8 (*α, β*-C-2′,6′, 2″,6″), 97.9 (*β*-C-1), 93.4 (*α*-C-1), 78.7 (*β*-C-3), 76.8 (*α*-C-3), 74.7 (*β*-C-5), 74.0 (*β*-C-2), 71.7 (*α*-C-5), 70.5 (*α*-C-2), 69.5 (*α, β*-C-4), 64.4 (*β*-C-6), 63.3 (*α*-C-6)

3,4,6-Tri-*O*-galloyl-D-glucose (**33**)

Yellowish amorphous powder

$C_{27}H_{24}O_{18}$

$[\alpha]_D^{20}$= -3.5 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.8), 278 (4.5) nm

IR ν_{max}: 3370, 1701, 1618, 1514, 1455, 1340, 1222 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 635.0889 $[M - H]^-$ (calcd for $C_{27}H_{23}O_{18}$ 635.0884)

^{1}H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 10

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 165.7 (*α*-C-7‴), 165.6 (*α, β*-C-7′), 165.4 (*β*-C-7″), 165.0 (*α, β*-C-7″), 145.2, 145.1, 145.0 (*α, β*-C-3′,5′, 3″,5″, 3‴,5‴), 138.2, 138.0, 137.9 (*α, β*-C-4′, 4″, 4‴), 121.0, 120.9, 120.1 (*α, β*-C-1′, 1″, 1‴), 109.3, 109.2 (2C) (*α, β*-C-2′,6′, 2″,6″, 2‴,6‴), 97.6 (*β*-C-1), 92.9 (*α*-C-1), 75.1 (*β*-C-3), 73.8 (*β*-C-2), 73.2 (*α*-C-3), 72.1 (*β*-C-5), 71.3 (*α*-C-2), 69.3 (*β*-C-4), 69.2 (*α*-C-4), 67.8 (*α*-C-5), 62.7 (*α, β*-C-6)

Chebulic acid (**34**)

Colorless syrup

$C_{14}H_{12}O_{11}$

$[\alpha]_D^{20}$= 20.5 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.6), 278 (4.2) nm

CD (MeOH): 217 (Δε -7.3), 235 (Δε 9.5), 279 (Δε 3.0)

IR ν_{max}: 3371, 1708, 1516, 1465, 1368 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 355.0302 $[M - H]^-$ (calcd for $C_{14}H_{11}O_{11}$ 355.0301)

^1H (400 MHz, DMSO-d_6) NMR: See Table 11

^{13}C (100 MHz, DMSO-d_6) NMR: See Table 12

6-*O*-Methyl chebulate (**35**)

Yellowish syrup

$C_{15}H_{14}O_{11}$

$[\alpha]_D^{20}$= 28.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (4.3), 284 (3.8) nm

IR ν_{max}: 3372, 1708, 1516, 1386 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 369.0458 $[M - H]^-$ (calcd for $C_{15}H_{13}O_{11}$ 369.0458)

^1H (400 MHz, CD_3OD) NMR: See Table 11

^{13}C (100 MHz, CD_3OD) NMR: See Table 12

7-*O*-Methyl chebulate (**36**)

Yellowish syrup

$C_{15}H_{14}O_{11}$

$[\alpha]_D^{20}$= 41.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 223 (4.3), 286 (3.9) nm

IR ν_{max}: 3381, 1707, 1516, 1465, 1367 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 369.0436 $[M - H]^-$ (calcd for $C_{15}H_{14}O_{11}$ 369.0458)

^1H (400 MHz, CD_3OD) NMR: See Table 11

^{13}C (100 MHz, CD_3OD) NMR: See Table 12

Neochebulic acid (**37**)

Colorless syrup

$C_{14}H_{12}O_{11}$

$[\alpha]_D^{20}$= 27.6 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 224 (4.4), 288 (4.0) nm

CD (MeOH): 214 (Δε -3.2), 233 (Δε 9.6), 281 (Δε 1.5)

IR ν_{max}: 3378, 1710, 1514, 1463, 1370 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 355.0305 $[M - H]^-$ (calcd for $C_{15}H_{14}O_{11}$ 355.0301)

^1H (400 MHz, DMSO-d_6) NMR: See Table 11

^{13}C (100 MHz, DMSO-d_6) NMR: See Table 12

m-Digallic acid and *p*-digallic acid (**38a** and **38b**)

White amorphous powder

$C_{14}H_{10}O_9$

UV (MeOH) λ_{max} (log ε): 284 (3.7) nm

IR ν_{max}: 3354, 1696, 1613, 1532, 1448, 1321, 1202 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 321.0228 $[M - H]^-$ (calcd for $C_{14}H_{10}O_9$ 321.0247)

^1H (400 MHz, CD_3OD) NMR: (*m*-digallic acid) δ 7.40 (1H, *d*, *J* = 1.8 Hz, H-2), 7.26 (1H, *d*, *J* = 1.8 Hz, H-6), 7.22 (2H, *s*, H-2', 6'), (*p*-digallic acid) δ 7.23 (6/7 H, *s*, H-2', 6') 7.11 (6/7 H, *s*, H-2, 6)

^{13}C (100 MHz, CD_3OD) NMR: (*m*-digallic acid) δ 169.7 (C-7), 166.3 (C-7'), 147.5 (C-3), 146.6 (C-3', 5'), 144.1 (C-4), 140.5 (C-4'), 140.1 (C-5), 122.4 (C-1), 120.5 (C-1'), 117.4 (C-6), 115.0 (C-2), 110.7 (C-2', 6'), (*p*-digallic acid) δ 169.9 (C-7), 164.6 (C-7'), 151.6 (C-3, 5), 146.5 (C-3', 5'), 140.3 (C-4'), 132.5 (C-4), 130.0 (C-1), 120.6 (C-1'), 110.9 (C-2', 6'), 110.0 (C-2, 6)

Ellagic acid (**39**)

Yellow amorphous powder

$C_{14}H_6O_8$

UV (MeOH) λ_{max} (log ε): 255 (4.7) and 366 (4.1) nm

IR ν_{max}: 3073, 1706, 1619, 1446, 1393, 1355, 1107 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 300.9983 $[M - H]^-$ (calcd for $C_{14}H_6O_8$ 300.9984)

^1H (300 MHz, DMSO-d_6) NMR: δ 7.45 (2H, *s*, H-5, 5')

^{13}C (75 MHz, DMSO-d_6) NMR: δ 159.1 (C-7, 7'), 148.1 (C-4, 4'), 139.6 (C-3, 3'), 136.3 (C-2, 2'), 112.3 (C-1, 1'), 110.2 (C-5, 5'), 107.6 (C-6, 6')

Brevifolincarboxylic acid (**40**)

White amorphous powder

$C_{13}H_8O_8$

$[\alpha]_D^{20}$= 21.6 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 278 (4.3) nm

IR ν_{max}: 3125, 1707, 1597, 1391, 1095, 980 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 291.0133 $[M - H]^-$ (calcd for $C_{13}H_7O_8$ 291.0141)

^{1}H (300 MHz, DMSO-d_6) NMR: δ 7.27 (1H, *s*, H-3'), 4.36 (1H, *dd*, *J* = 2.0, 8.0 Hz, H-4), 2.91 (1H, *dd*, *J* = 8.0, 18.5 Hz, H-5a), 2.52 (1H, *m*, H-5b)

^{13}C (75 MHz, DMSO-d_6) NMR: δ 193.6 (C-1), 173.4 (C-6), 160.3 (C-7'), 149.4 (C-2), 145.2 (C-3), 144.5 (C-6'), 140.6 (C-4'), 139.4 (C-5'), 115.2 (C-2'), 113.0 (C-1'), 108.0 (C-3'), 41.2 (C-4), 37.5 (C-5)

Phyllanemblinin F (**41**)

Light brown amorphous powder

$C_{27}H_{26}O_{20}$

$[\alpha]_D^{20}$= 6.8 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.7), 281 (4.3) nm

IR ν_{max}: 3394, 1722, 1616, 1329, 1206, 1064 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 669.0923 $[M - H]^-$ (calcd for $C_{27}H_{25}O_{20}$ 669.0916)

^1H (300 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 7.18 (2H, *s*, Gal-H-2', 6'), 7.13 (1H, *s*, Neocheb-H-2''), 5.69 (1H, *d*, *J* = 7.8 Hz, H-1), 5.31 (1H, *d*, *J* = 0.9 Hz, Neocheb-H-2'), 4.64 (1H, *dd*, *J* = 1.6, 12.0 Hz, H-6a), 4.33 (1H, *dd*, *J* = 4.2, 12.0 Hz, H-6b), 3.92 (1H, *d*, *J* = 8.4 Hz, Neocheb-H-3'), 3.80–3.60 (4H, m H-2, H-3, H-4, H-5), 3.19 (1H, *ddd*, *J* = 4.2, 8.4, 11.1 Hz, Neocheb-H-4'), 2.90 (1H, *dd*, *J* = 11.1, 17.1 Hz, Neocheb-H-5'a), 2.43 (1H, *dd*, *J* = 4.2, 17.1 Hz, Neocheb-H-5'b)

^{13}C (75 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 174.4 (Neocheb-C-1') , 174.3 (Neocheb-C-6'), 171.7 (Neocheb-C-7'), 166.2 (Neocheb-C-7''), 165.3 (Gal-C-7'), 146.3 (Neocheb-C-3''), 146.1 (Gal-C-3', 5'), 143.5 (Neocheb-C-5''), 139.7 (Neocheb-C-4''), 139.6 (Gal-C-4'), 120.5 (Gal-C-1'), 117.5 (Neocheb-C-6''), 116.5 (Neocheb-C-1''), 110.5 (Gal-C-2', 6'), 109.3 (Neocheb-C-2''), 95.6 (C-1), 78.2 (Neocheb-C-2'), 77.4 (C-5), 75.8 (C-3), 73.5 (C-2), 70.4 (C-4), 64.2 (C-6), 45.7 (Neocheb-C-4'), 36.9 (Neocheb-C-3'), 34.9 (Neocheb-C-5')

Corilagin (**42**)

Yellowish amorphous powder

$C_{27}H_{22}O_{18}$

$[\alpha]_D^{20}$= -147.7 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.8), 272 (4.5) nm

CD (MeOH): 237 ($\Delta\varepsilon$ -27.8), 263 ($\Delta\varepsilon$ 8.1), 284 ($\Delta\varepsilon$ -19.4)

IR ν_{max}: 3411, 1712, 1614, 1328, 1207, 1034 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 663.0720 $[M - H]^-$ (calcd for $C_{27}H_{21}O_{18}$ 663.0728)

^1H (400 MHz, CD_3OD) NMR: See Table 13

Table 13^{13}C (100 MHz, CD_3OD) NMR: See Table 13

Tercatain (**43**)

Light brown amorphous powder

$C_{34}H_{26}O_{22}$

$[\alpha]_D^{20}$= -29.9 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 219 (4.8), 278 (4.4) nm

IR ν_{max}: 3332, 1708, 1613, 1452, 1341, 1204, 1032, 970 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 785.0840 $[M - H]^-$ (calcd for $C_{34}H_{25}O_{22}$ 785.0837)

^1H (400 MHz, acetone-d_6) NMR: See Table 13

^{13}C (100 MHz, acetone-d_6) NMR: See Table 13

Gemin D (**44**)

Light brown amorphous powder

$C_{27}H_{22}O_{18}$

$[\alpha]_D^{20}$= 44.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.6), 266 (4.3) nm

IR ν_{max}: 3387, 1707, 1616, 1449, 1329, 1235, 1029 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 633.0722 [M – H]$^-$ (calcd for $C_{27}H_{21}O_{18}$ 633.0728)

^{1}H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 14

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 168.6 (*α*-HHDP-C-7''), 168.5 (*β*-HHDP-C-7''), 168.0 (*α*-HHDP-C-7'), 167.9 (*β*-HHDP-C-7'), 167.5 (*α*-Gal-C-7'), 167.3 (*β*-Gal-C-7'), 145.7 (α, β-Gal-C-3', 5') 145.1 (*α*, *β*-HHDP-C-4', 4''), 144.3 (*α*, *β*-HHDP-C-2', 2''), 138.9 (*α*, *β*-Gal-C-4'), 136.4 (*α*, *β*-HHDP-C-3'), 136.3 (*α*, *β*-HHDP-C-3''), 126.3, 126.0 (*α*, *β*-HHDP-C-3', 3''), 121.1 (*α*-Gal-C-1'), 121.0 (*β*-Gal-C-1'), 115.8 (*α*, *β*-HHDP-C-1''), 115.7 (*α*, *β*-HHDP-C-1'), 110.1 (*α*, *β*-Gal-C-2', 6'), 107.9 (*α*, *β*-HHDP-C-5''), 107.8 (*α*, *β*-HHDP-C-5'), 98.6 (*β*-C-1), 93.7 (*α*-C-1), 76.0 (*β*-C-3), 74.5 (*α*-C-3), 74.3 (*β*-C-2), 71.7 (*α*-C-2, *β*-C-5) 71.3 (*α*, *β*-C-4), 67.1 (*α*-C-5), 63.8 (*α*, *β*-C-6)

Tellimagrandin I (**45**)

Light brown amorphous powder

$C_{34}H_{26}O_{22}$

$[\alpha]_D^{20}$= 40.9 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.8), 272 (4.5) nm

IR ν_{max}: 3395, 1708, 1616, 1327, 1215, 1034 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 785.0840 [M – H]$^-$ (calcd for $C_{34}H_{25}O_{22}$ 785.0837)

^{1}H (400 MHz, CD_3OD) NMR: See Table 14

^{13}C (100 MHz, CD_3OD) NMR: δ 168.1 (*α*-HHDP-C-7''), 168.0 (*β*-HHDP-C-7''), 167.6 (*α*, *β*-HHDP-C-7'), 166.5 ((*α*-Gal-C-7''), 166.3 (*β*-Gal-C-7''), 166.0 (*α*-Gal-C-7'), 165.5 (*β*-Gal-C-7'), 145.8 (α, β-Gal-C-3', 5'), 145.6 (α, β-Gal-C-3'', 5''), 145.1

(*β*-HHDP-C-4"), 145.0 (*α*-HHDP-C-4"), 144.3 (*β*-HHDP-C-4'), 144.2 (*α*-HHDP-C-4"), 139.0 (*α*, *β*-Gal-C-4'), 138.8 (*α*, *β*-Gal-C-4"), 136.3 (*α*, *β*-HHDP-C-3', 3"), 126.6, 126.1 (*α*, *β*-HHDP-C-3', 3"), 120.9 (*α*, *β*-Gal-C-1') , 120.8 (*α*, *β*-Gal-C-1"), 115.6 (*α*, *β*-HHDP-C-1'), 115.4 (*α*, *β*-HHDP-C-1"), 110.0 (*α*, *β*-Gal-C-2', 6', *α*, *β*-Gal-C-2", 6"), 108.0 (*α*, *β*-HHDP-C-5"), 107.7 (*α*, *β*-HHDP-C-5') 96.7 (*β*-C-1), 91.2 (*α*-C-1), 74.1 (*β*-C-2), 73.5 (*β*-C-3), 73.0 (*α*-C-2), 72.0 (*β*-C-1), 71.1 (*α*-C-3, *α*, *β*-C-4), 67.2 (*α*-C-5), 63.5 (*α*, *β*-C-6)

Punicacortein C (**46**)

Yellowish amorphous powder

$C_{48}H_{28}O_{30}$

$[\alpha]_D^{20}$= -28.3 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 216 (5.0), 260 (4.9), 376 (4.2) nm

IR ν_{max}: 3393, 1707, 1606, 1319, 1178, 1058 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1083.0581 $[M - H]^-$ (calcd for $C_{43}H_{27}O_{30}$ 1083.0587)

^{1}H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: *δ* 7.28 (1H, *s*, Gallagyl-H-2"), 6.70 (1H, *s*, Gallagyl-H-2'), 6.35 (1H, *s*, HHDP-H-5"), 5.41 (1H, *d*, *J* = 4.7 Hz, H-1), 5.13 (1H, *dd*, *J* = 2.2, 5.0 Hz, H-3), 4.87 (1H, *dd*, *J* = 2.2, 4.7 Hz, H-2), 4.23 (1H, *dd*, *J* = 1.9, 5.2 Hz, H-4), 3.81 (2H, *m*, H-6), 2.30 (1H, *d*, *J* = 9.3 Hz, H-5)

^{13}C (400 MHz, acetone-d_6:D_2O = 9:1) NMR: *δ* 168.8 (HHDP-C-7"), 168.0 (Gallagyl-C-7'), 167.9 (Gallagyl-C-7"), 165.3 (HHDP-C-7'), 159.8, 158.9 (Gallagyl-C-7"', Gallagyl-C-7""), 147.6, 147.0, 146.0 (HHDP-C-4"), 145.8 (Gallagyl-C-3'), 145.4 (Gallagyl-C-3"), 145.1, 144.8, 144.5 (HHDP-C-4'), 143.8,

143.8, 139.5, 139.0, 138.3, 138.3 (Gallagyl-C-4"), 138.1, 136.8 (Gallagyl-C-4'), 136.3, 134.9 (HHDP-C-3"), 127.2, 125.0, 124.7, 122.7, 120.4, 120.3, 118.2 (Gallagyl C-1"), 116.6 (HHDP-C-5'), 116.2 (HHDP-C-1"), 115.9 (Gallagyl-C-1'), 115.3, 114.9, 114.4 (Gallagyl C-2"), 112.5, 110.6, 110.1 (Gallagyl-C-2'), 104.8 (HHDP-C-5"), 74.5 (C-2), 73.4 (C-4), 72.8 (C-5), 68.0 (C-1), 67.7 (C-6), 67.4 (C-3)

Punicacortein D (**47**)

Light brown amorphous powder

$C_{48}H_{28}O_{30}$

$[\alpha]_D^{20}$= -70.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 215 (4.9), 259 (4.8), 375 (4.2) nm

IR ν_{max}: 3394, 1709, 1605, 1320, 1178, 1059 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1083.0580 $[M - H]^-$ (calcd for $C_{43}H_{27}O_{30}$ 1083.0587)

^1H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 7.27 (1H, *s*, Gallagyl-H-2"), 6.80 (1H, *s*, Gallagyl-H-2'), 6.34 (1H, *s*, HHDP-H-5"), 4.79 (1H, *s*, H-3), 4.78 (1H, *s*, H-2), 4.63 (1H, *s*, H-1), 4.45 (1H, br *s*, H-4), 3.82 (1H, *dd*, *J* = 9.5, 12.0 Hz, H-6a), 3.68 (1H, overlapped, H-6b) 2.36 (1H, *d*, *J* = 9.5 Hz, H-5)

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 168.6 (HHDP-C-7"), 168.5 (Gallagyl-C-7'), 168.1 (Gallagyl-C-7"), 166.1 (HHDP-C-7'), 159.1, 159.0 (Gallagyl-C-7'", Gallagyl-C-7""), 148.2, 147.3, 146.5 (HHDP-C-4"), 145.8 (Gallagyl-C-3'), 145.4 (Gallagyl-C-3"), 145.1, 145.1, 144.4 (HHDP-C-4'), 144.1, 143.4, 139.8, 139.1, 138.3, 138.0 (Gallagyl-C-4"), 137.7, 136.8 (Gallagyl-C-4'), 136.6, 134.9 (HHDP-C-3"), 127.5, 124.7, 124.6, 122.7, 120.4, 120.1, 118.5

(Gallagyl C-1"), 116.1 (HHDP-C-5'), 116.0 (HHDP-C-1"), 116.0 (Gallagyl-C-1'), 115.3, 115.0, 114.1 (Gallagyl C-2"), 112.3, 110.1, 108.7 (Gallagyl-C-2'), 105.0 (HHDP-C-5"), 79.2 (C-2), 74.9 (C-1), 72.5 (C-5), 69.8 (C-3), 67.9 (C-6), 65.2 (C-4)

Punicalagin (**48**)

Light brown amorphous powder

$C_{48}H_{28}O_{30}$

$[\alpha]_D^{20}$= -112.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.9), 259 (4.9), 380 (4.2) nm

IR ν_{max}: 3411, 1725, 1606, 1328, 1175, 1055 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1083.0574 $[M - H]^-$ (calcd for $C_{48}H_{27}O_{30}$ 1083.0587)

^1H (600 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 15

^{13}C (150 MHz, acetone-d_6:D_2O = 9:1) NMR: δ 169.4, 169.3 (*α*, *β*-Gallagyl-C-7') (*α*, *β*-HHDP-C-7"), 168.8, 168.6 (*α*, *β*-HHDP-C-7'), 168.5, 168.3 (*α*, *β*-HHDP-C-7"), 168.0, 167.9 (*α*, *β*-Gallagyl-C-7"), 158.2, 157.7 (*α*, *β*-Gallagyl-C-7"', Gallagyl-C-7""), 148.4, 147.9, 147.8, 147.8, 145.6, 145.5, 145.4, 145.2, 145.1, 145.1, 145.0, 144.8, 144.7, 144.4, 144.3, 144.3, 144.2, 144.1, 144.0, 139.0, 138.7, 138.3, 138.3, 137.6, 137.6 (*α*, *β*-Gallagyl-C-4"), 137.4, 137.0, 136.9, 136.8, 136.6 (*β*-HHDP-C-3'), 136.3, 136.2, 136.2, 136.1, 126.9, 126.8, 126.8, 126.6, 125.7, 125.1, 124.5, 124.2, 123.8, 122.0, 121.9, 118.0, 117.9 (*α*, *β*-Gallagyl-C-1"), 115.0, 114.8, 114.8, 114.6, 114.5, 114.5, 114.3, 114.2, 114.0, 111.8, 111.7 (*α*, *β*-Gallagyl-C-2"), 111.4, 111.0, 111.0, 109.6 (*β*-HHDP-C-5"), 108.9, 107.6, 107.5, 107.3, 94.4 (*β*-C-1), 90.3 (*α*-C-1), 79.1 (*β*-C-3), 76.9 (*β*-C-2), 76.6 (*α*-C-3), 74.5 (*α*-C-2), 72.6 (*β*-C-5), 71.2

(*α*-C-4), 71.0 (*β*-C-4), 67.0 (*α*-C-5), 64.4 (*α*-C-6), 64.3 (*β*-C-6)

Terflavin A (**49**)

Light brown amorphous powder

$C_{48}H_{30}O_{30}$

$[\alpha]_D^{20}$= 203.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.8), 257 (4.8), 371 (4.1) nm

IR ν_{max}: 3393, 1707, 1515, 1329, 1187, 1051 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1085.0734 $[M - H]^-$ (calcd for $C_{48}H_{29}O_{30}$ 1085.0744)

^1H (400 MHz, acetone-d_6:D_2O = 9:1) NMR: See Table 15

^{13}C (100 MHz, acetone-d_6:D_2O = 9:1) NMR: *δ* 169.5 (*β*-HHDP-C-7'), 169.4 (*α*-HHDP-C-7'), 169.3 (*α*-HHDP-C-7''), 169.1 (*β*-HHDP-C-7'), 166.9 (*α*-Gal-C-7'), 166.8 (*β*-Gal-C-7'), 166.00 (*β*-Flavogallonyl-C-7'), 166.0 (*α*-Flavogallonyl-C-7'), 160.4 (*α*-Flavogallonyl-C-7'''), 160.3 (*β*-Flavogallonyl-C-7'''), 158.8, 148.8 (*α*-Flavogallonyl-C-4'''), 148.7 (*β*-Flavogallonyl-C-4'''), 147.1, 147.0, 146.1 (*α*-Gal-C-3', 5'), 146.1 (*β*-Gal-C-3', 5'), 145.6 (*α*-HHDP-C-4''), 145.5, 145.5 (*α*-HHDP-C-4'), 145.4, 144.9, 144.9 (*β*-HHDP-C-4''), 144.7, 144.6 (*β*-HHDP-C-4'), 139.6, 139.5, 139.4, 139.3, 139.1 (*α*-Gal-C-4', *β*-Gal-C-7'), 138.0, 138.0, 137.5, 137.5, 136.8 (*β*-HHDP-C-3'), 136.8 (*α*-HHDP-C-4'), 136.8 (*α*-HHDP-C-3'), 136.7 (*β*-HHDP-C-3''), 127.3, 127.1, 127.0, 124.9, 124.7, 121.7, 121.6 (*α*-Gal-C-1'), 121.6, 121.5 (*β*-Gal-C-1'), 118.5, 115.0 (*β*-HHDP-C-1''), 114.9 (*α*-HHDP-C-1''), 114.9 (*β*-HHDP-C-1'), 114.8 (*α*-HHDP-C-1'), 114.5, 114.5, 114.4, 114.3, 111.8 (*α*-Flavogallonyl-C-5'''), 111.7 (*β*-Flavogallonyl-C-5'''), 111.5 (*α*-Flavogallonyl-C-2''), 111.5 (*β*-

Flavogallonyl-C-2'), 110.3 (*α*-Gal-C-2', 6'), 110.2, 110.2 (*β*-Gal-C-2', 6'), 109.4, 109.3, 108.4 (*α*-HHDP-C-5'), 108.2 (*β*-HHDP-C-5'), 108.2 (*α*-HHDP-C-5''), 107.8 (*β*-HHDP-C-5''), 95.5 (*β*-C-1), 91.8 (*α*-C-1), 78.2 (*β*-C-2), 77.9 (*β*-C-3), 75.6 (*α*-C-2), 75.5 (*α*-C-3), 73.5 (*β*-C-5), 68.8 (*α*-C-5), 68.5 (*α*-C-4), 68.1 (*β*-C-4), 63.0 (*α*-C-6), 62.9 (*β*-C-6)

Arjungenin (**50**)

White amorphous powder

$C_{30}H_{48}O_6$

$[\alpha]_D^{20}$= 52.2 (c 0.10, MeOH)

IR ν_{max}: 3394, 2938, 1693, 1462, 1385, 1270, 1042 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 503.3382 $[M - H]^-$ (calcd for $C_{30}H_{47}O_6$ 503.3373)

^1H (600 MHz, pyridine-d_5) NMR: See Table 16

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 17

23-*O*-Neochebuloylarjungenin 28-*O*-*β*-D-glucopyranosyl ester (**51**)

Brown amorphous powder

$C_{50}H_{68}O_{21}$

$[\alpha]_D^{20}$= -5.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 218 (4.3), 284 (3.8) nm

IR ν_{max}: 3392, 2939, 1694, 1515, 1464, 1386, 1042 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 1003.4206 $[M - H]^-$ (calcd for $C_{50}H_{67}O_{21}$ 1003.4175)

^1H (800 MHz, pyridine-d_5) NMR: See Table 16

^{13}C (200 MHz, pyridine-d_5) NMR: See Table 17

23-*O*-4′-*epi*-Neochebuloylarjungenin (**52**)

Light brown amorphous powder

$C_{44}H_{58}O_{16}$

$[\alpha]_D^{20}$= 10.3 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.4), 284 (4.0) nm

IR ν_{max}: 3384, 2932, 1701, 1515, 1427, 1050 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 841.3647 $[M - H]^-$ (calcd for $C_{44}H_{58}O_{16}$ 841.3647)

^1H (800 MHz, pyridine-d_5) NMR: See Table 16

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 17

23-*O*-Galloylarjunic acid (**53**)

White amorphous powder

$C_{37}H_{52}O_{10}$

$[\alpha]_D^{20}$= 29.1 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 216 (4.6), 278 (4.2) nm

IR ν_{max}: 3359, 2942, 1695, 1462, 1384, 1318, 1227, 1024 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 655.3484 $[M - H]^-$ (calcd for $C_{37}H_{51}O_{10}$ 655.3482)

^1H (400 MHz, pyridine-d_5) NMR: See Table 18

^{13}C (100 MHz, pyridine-d_5) NMR: See Table 18

Quercotriterpenoside I (**54**)

Brown amorphous powder

$C_{43}H_{62}O_{15}$

$[\alpha]_D^{20}$= 8.8 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.4), 284 (4.0) nm

IR ν_{max}: 3389, 2934, 1738, 1462, 1387, 1061 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 817.4030 $[M - H]^-$ (calcd for $C_{43}H_{61}O_{15}$ 817.4010)

^1H (600 MHz, pyridine-d_5) NMR: See Table 18

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 18

Arjunglucoside I (**55**)

White amorphous powder

$C_{36}H_{58}O_{11}$

$[\alpha]_D^{20}$= 14.7 (c 0.10, MeOH)

IR ν_{max}: 3387, 2937, 1693, 1463, 1327, 1233, 1040 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 711.3958 $[M + HCOOH - H]^-$ (calcd for $C_{37}H_{59}O_{13}$

711.3956)

^{1}H (500 MHz, pyridine-d_5) NMR: δ 6.39 (1H, *d*, *J* = 8.2 Hz, H-1′), 5.51 (1H, br *s*, H-12), 4.46 (1H, *dd*, *J* = 2.1, 11.8 Hz, H-6′a), 4.41 (1H, *dd*, *J* = 5.0, 11.8 Hz, H-6′b), 4.38 (1H, *t*, *J* = 9.3 Hz, H-4′), 4.31 (1H, *t*, *J* = 8.9 Hz, H-3′), 4.26 (1H, *m*, H-2) 4.22 (1H, *t*, *J* = 8.7 Hz, H-2′), 4.20 (1H, *d*, *J* = 9.9 Hz, H-3), 4.20 (1H, *d*, *J* = 10.5 Hz, H-23a), 4.04 (1H, m, H-5′), 3.73 (1H, *d*, *J* = 10.5, H-23b), 3.57 (1H, br *s*, H-19), 3.53 (1H, br *s*, H-18), 2.81 (1H, *td*, *J* = 13.2, 3.1 Hz, H-16a), 2.37 (1H, *td*, *J* = 13.8, 2.4 Hz, H-15a), 2.30 (1H, *dd*, *J* = 4.2, 12.4 Hz, H-1a), 2.11 (1H, m, H-16b), 2.10 (1H, *m*, H-9), 2.10 (2H, *m*, H-11), 2.06 (1H, *m*, H-22a), 2.05 (1H, *m*, H-21a), 1.95 (1H, m, H-22b), 1.85 (1H, br *d*, *J* = 12.0 Hz, H-5), 1.72 (1H, *m*, H-6a), 1.71 (1H, *m*, H-7a), 1.56 (3H, *s*, H-27), 1.49 (1H, *m*, H-6b), 1.46 (1H, *m*, H-7b), 1.38 (1H, *t*, *J* = 11.8 Hz, H-1b), 1.21 (1H, *m*, H-15b), 1.21 (3H, *s*, H-25), 1.15 (3H, *s*, H-26), 1.15 (3H, *s*, H-29), 1.09 (3H, *s*, H-24), 1.05 (1H, *m*, H-6b), 0.99 (3H, *s*, H-30)

^{13}C (125 MHz, pyridine-d_5) NMR: δ 177.8 (C-28), 144.9 (C-13), 124.3 (C-12), 96.4 (C-1′), 81.5 (C-19), 79.8 (C-5′), 79.4 (C-3′), 78.8 (C-3), 74.6 (C-2′), 71.6 (C-4′), 69.4 (C-2), 67.0 (C-23), 62.7 (C-6′), 49.0 (C-9), 48.6 (C-5), 48.0 (C-1), 47.0 (C-17), 45.1 (C-18), 44.2 (C-4), 42.7 (C-14), 40.8 (C-8), 39.1 (C-10), 36.0 (C-20), 33.6 (C-22), 33.4 (C-7), 29.5 (C-15), 29.4 (C-21), 29.3 (C-29), 28.5 (C-16), 25.4 (C-27), 25.2 (C-30), 24.8 (C-11), 19.3 (C-6), 18.2 (C-25), 17.9 (C-26), 14.8 (C-24)

Terminolic acid (**56**)

White amorphous powder

$C_{30}H_{48}O_6$

$[\alpha]_D^{20}$= 44.7 (c 0.10, MeOH)

IR ν_{max}: 3411, 2938, 1693, 1463, 1386, 1044 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 503.3384 $[M - H]^-$ (calcd for $C_{30}H_{47}O_6$ 503.3373)

^1H (600 MHz, pyridine-d_5) NMR: See Table 19

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 19

23-*O*-Galloylterminolic acid 28-*O*-*β*-D-glucopyranosyl ester (**57**)

Light brown amorphous powder

$C_{43}H_{62}O_{15}$

$[\alpha]_D^{20}$= 4.7 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 216 (4.5), 278 (4.1) nm

IR ν_{max}: 3412, 2939, 1706, 1462, 1365, 1231, 1062 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 817.4025 $[M - H]^-$ (calcd for $C_{43}H_{61}O_{15}$ 817.4010)

^1H (600 MHz, pyridine-d_5) NMR: See Table 19Table 19^{13}C (150 MHz, pyridine-d_5) NMR: See Table 19

Arjunolic acid (**58**)

White amorphous powder

$C_{30}H_{48}O_5$

$[\alpha]_D^{20}$= 47.8 (c 0.10, MeOH)

IR ν_{max}: 3391, 2938, 1693, 1515, 1463, 1394, 1044 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 487.3436 $[M - H]^-$ (calcd for $C_{30}H_{47}O_5$ 487.3423)

^1H (400 MHz, pyridine-d_5) NMR: See Table 20

^{13}C (100 MHz, pyridine-d_5) NMR: See Table 20

Arjunglucoside II (**59**)

White amorphous powder

$C_{36}H_{58}O_{10}$

$[\alpha]_D^{20}$ = 51.8 (c 0.10, MeOH)

IR ν_{max}: 3365, 2939, 1726, 1463, 1061 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 695.4022 $[M + HCOOH - H]^-$ (calcd for $C_{37}H_{59}O_{12}$ 695.4007)

^1H (400 MHz, pyridine-d_5) NMR: See Table 20

^{13}C (100 MHz, pyridine-d_5) NMR: See Table 20

23-O-Galloylarjunolic acid (**60**)

White amorphous powder

$C_{37}H_{52}O_9$

$[\alpha]_D^{20}$ = 33.0 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 216 (4.5), 276 (4.0) nm

IR ν_{max}: 3401, 2923, 1695, 1684, 1201 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 639.3535 $[M– H]^-$ (calcd for $C_{37}H_{51}O_9$ 639.3535)

^1H (600 MHz, CD_3OD) NMR: See Table 21

^{13}C (150 MHz, CD_3OD) NMR: See Table 21

23-*O*-Galloylarjunolic acid 28-*O*-*β*-D-glucopyranosyl ester (**61**)

Light brown amorphous powder

$C_{43}H_{62}O_{14}$

$[\alpha]_D^{20}$= 25.2 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 220 (4.4), 284 (4.0) nm

IR ν_{max}: 3418, 2943, 1695, 1613, 1348, 1236, 1031 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 801.4066 $[M – H]^-$ (calcd for $C_{43}H_{61}O_{14}$ 801.4061)

^1H (600 MHz, pyridine-d_5) NMR: See Table 21

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 21

Arjunic acid (**62**)

White amorphous powder

$C_{30}H_{48}O_5$

$[\alpha]_D^{20}$= 32.0 (c 0.10, MeOH)

IR ν_{max}: 3404, 2935, 1692, 1515, 1463, 1296, 1042 cm^{-1}

HRMS (ESI-QTOF) *m/z*: 487.3432 $[M – H]^-$ (calcd for $C_{30}H_{47}O_5$ 487.3423)

^{1}H (600 MHz, pyridine-d_5) NMR: See Table 22

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 22

Arjunetin (**63**)

White amorphous powder

$C_{36}H_{58}O_{10}$

$[\alpha]_D^{20}$= 32.4 (c 0.10, MeOH)

IR ν_{max}: 3394, 2938, 1739, 1515, 1463, 1387, 1061 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 695.4018 [M + HCOOH – H]$^-$ (calcd for $C_{37}H_{59}O_{12}$ 695.4007)

^{1}H (600 MHz, pyridine-d_5) NMR: See Table 22

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 22

Crataegioside (**64**)

White amorphous powder

$C_{36}H_{58}O_{10}$

$[\alpha]_D^{20}$= 33.8 (c 0.10, MeOH)

IR ν_{max}: 3394, 2936, 1707, 1515, 1463, 1348, 1200, 1063 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 695.4011 [M + HCOOH – H]$^-$ (calcd for $C_{37}H_{59}O_{12}$ 695.4007)

^{1}H (600 MHz, pyridine-d_5) NMR: δ 6.42 (1H, *d*, *J* = 8.1, H-1′), 5.55 (1H, br *t*, *J* = 4.0 Hz, H-12), 4.47 (1H, *m*, H-6′a), 4.44 (1H, *m*, H-6′b), 4.41 (1H, *t*, *J* = 9.2 Hz, H-4′), 4.32 (1H, *t*, *J* = 9.1, H-3′), 4.24 (1H, *t*, *J* = 8.6 Hz, H-2′), 4.23 (1H, *m*, H-3), 4.19 (1H, *d*, *J* = 10.3 Hz, H-23a), 4.06 (1H, *m*, H-5′), 3.74 (1H, *d*, *J* = 10.3 Hz, H-23b), 3.60 (1H, br *s*, H-19), 3.56 (1H, br *s*, H-18), 2.84 (1H, br *t*, *J* = 11.9 Hz, H-16a), 2.40 (1H, *t*, *J* = 12.9 Hz, H-15a), 2.14 (1H, d, *J*=12.2 Hz, H-16b), 2.06 (1H, *m*, H-7a), 2.06 (1H, *m*, H-21a), 2.05 (2H, *m*, H-11), 1.97 (1H, *m*, H-9), 1.97 (1H, *m*, H-22a), 1.95 (1H, *m*, H-2a), 1.90 (1H, *m*, H-2b), 1.68 (1H, *m*, H-22b), 1.67 (1H, *m*, H-6a), 1.60 (3H, *s*, H-27), 1.58 (1H, *m*, H-1a), 1.57 (1H, *m*, H-5), 1.47 (1H, *m*, H-7b), 1.46 (1H, *m*, H-6), 1.25 (1H, *m*, H-15b), 1.22 (3H, *s*, H-26), 1.16 (3H, *s*, H-29), 1.09 (3H, *s*, H-24), 1.07 (1H, *m*, H-1b), 1.06 (1H, *m*, H-21b), 1.05 (3H, *s*, H-25), 1.00 (3H, *s*, H-30)

^{13}C (150 MHz, pyridine-d_5) NMR: δ 177.8 (C-28), 144.8 (C-13), 124.3 (C-12), 96.4 (Glc C-1′), 96.4 (Glc C-1′), 81.5 (C-19), 79.8 (Glc C-5′), 79.4 (Glc C-3′), 74.6 (Glc C-2′), 73.9 (C-3), 71.5 (Glc C-4′), 68.3 (C-23), 62.6 (Glc C-6′), 49.1 (C-5), 48.9 (C-9), 47.0 (C-17), 45.1 (C-18), 43.4 (C-4), 42.6 (C-14), 40.7 (C-8), 39.1 (C-1), 37.9 (C-10), 36.0 (C-20), 33.5 (C-22), 33.4 (C-7), 29.5 (C-15), 29.4 (C-21), 29.2 (C-29), 28.5 (C-16), 28.2 (C-2), 25.4 (C-27), 25.1 (C-30), 24.7 (C-11), 19.3 (C-6), 18.1 (C-26), 16.4 (C-25), 13.5 (C-24)

Pinfaenoic acid 28-*O*-*β*-D-glucopyranosyl ester (**65**)

White amorphous powder

$C_{36}H_{56}O_{10}$

$[\alpha]_D^{20}$= 130.5 (c 0.10, MeOH)

IR ν_{max}: 3387, 2935, 1738, 1644, 1515, 1463, 1385, 1068 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 693.3864 $[M + HCOOH - H]^-$ (calcd for $C_{37}H_{57}O_{12}$ 693.3850)

^1H (600 MHz, pyridine-d_5) NMR: See Table 23

^{13}C (150 MHz, pyridine-d_5) NMR: See Table 23

23-*O*-Galloylpinfaenoic acid 28-*O*-*β*-D-glucopyranosyl ester (**66**)

Yellowish amorphous powder

$C_{43}H_{60}O_{14}$

$[\alpha]_D^{20}$ = 21.4 (c 0.10, MeOH)

UV (MeOH) λ_{max} (log ε): 217 (4.8), 282 (4.8) nm

IR ν_{max}: 3365, 2937, 1739, 1691, 1516, 1462, 1342, 1233, 1041 cm^{-1}

HRMS (ESI-QTOF) *m*/*z*: 799.3889 $[M - H]^-$ (calcd for $C_{43}H_{59}O_{14}$ 693.3850)

^1H (800 MHz, pyridine-d_5) NMR: See Table 23

^{13}C (200 MHz, pyridine-d_5) NMR: See Table 23

Table 5. ^{1}H NMR data for compounds **19**-**21**

	19[a]	**20**[b]	**21**[b]
β-D-Glucose			
1	5.79, *d* (8.1)	5.72, *d* (8.2)	6.05, *d* (8.2)
2	3.67, *dd* (8.5, 9.0)	3.64, *t* (8.6)	4.02, *m*
3	3.90, *dd*, (9.0, 9.5)	3.85, *t* (9.3)	5.67, *t* (9.5)
4	5.07, *t* (9.8)	4.99, *t* (9.7)	5.54, *t* (9.8)
5	3.79, *m*	3.78, *m*	4.38, *ddd* (1.8, 3.9, 12.6)
6a	3.77, *dd* (2.0, 12.5)	3.76, *dd* (2.0 12.4)	4.56, *dd* (1.5, 12.6)
6b	3.63, *dd* (6.0, 12.5)	3.58, *dd* (5.8, 12.4)	4.45, *dd* (3.9, 12.6)
1-*O*-Galloyl			
2′, 6′	7.19, *s*	7.16, *s*	7.19, *s*
4-*O*-Neochebuloyl			
2′	5.70, *d* (1.5)	5.68, *d* (0.9)	4.81, *s*
3′	3.88, *dd* (1.5, 10.5)	3.78, *dd* (0.9, 10.5)	3.78, *d* (11.2)
4′	3.14, *ddd* (3.0, 10.5, 12.0)	3.07, *td* (11.3, 3.2)	3.20, *ddd* (5.2, 9.1, 11.2)
5′a	3.07, *dd* (12.0, 17.0)	2.97, *dd* (11.6, 17.2)	2.80, *dd* (9.1, 17.5)
5′b	2.37, *dd* (3.0, 17.0)	2.28, *dd* (3.2, 17.2)	2.43, *dd* (5.2, 17.5)
2″	7.13, *s*	7.09, *s*	7.06, *s*
		1′-*O*-Me	1′-*O*-Me
		3.53, *s*	3.32, *s*
			6′-*O*-Me
			3.34, *s*
			3-*O*-Galloyl
2″, 6″			7.16, *s*
			6-*O*-Galloyl
2‴, 6‴			7.17, *s*

^{1}H NMR data were measured at [a]500 and [b]600 MHz in acetone-d_6:D_2O = 9:1, respectively.

Table 6. ^{13}C NMR data for compounds **19**-**21**

Position	**19**	**20**	**21**
β-D-Glucose			
1	95.5	95.2	95.2
2	74.5	73.9	72.8

3	75.5	74.9	75.9
4	72.2	71.9	69.9
5	76.6	76.3	73.4
6	61.8	61.5	62.8
1-*O*-Galloyl			
1′	121.1	120.2	120.7
2′, 6′	110.4	110.1	110.5
3′, 5′	146.1	146.0	146.1
4′	139.4	139.5	139.5
7′	165.5	165.9	165.2
4-*O*-Neochebuloyl			
1′	171.1	170.6	169.7
2′	77.3	77.4	77.5
3′	37.3	36.8	36.4
4′	45.9	45.7	44.9
5′	34.9	34.9	34.6
6′	174.5	174.5	172.2
7′	173.9	174.3	173.5
1″	117.1	115.8	116.9
2″	109.4	109.2	109.4
3″	146.0	146.2	145.9
4″	139.0	139.4	139.1
5″	143.5	143.3	143.5
6″	117.6	116.9	116.6
7″	164.0	164.4	163.4
		1′-*O*-Me	1′-*O*-Me
		52.9	52.8
			6′-*O*-Me
			51.9
			3-*O*-Galloyl
1″			121.2
2″, 6″			110.4
3″, 5″			146.0
4″			139.1
7″			166.3
			6-*O*-Galloyl
1‴			121.7
2‴, 6‴			110.1
3‴, 5‴			146.0
4‴			139.1
7‴			166.5

^{13}C NMR data were measured at [a]125 and [b]150 MHz in acetone-d_6:D_2O = 9:1, respectively.

Table 7. ^{1}H NMR data for compounds **22-24**

	22[a]	**23**[b]	**24**[b]
β-D-Glucose			
1	6.28, *d* (3.0)	6.16, *d* (4.1)	6.22, *d* (3.9)
2	4.08, *m*	4.20, br *d* (4.1)	4.22, br *d* (3.9)
3	4.90, *m*	4.90, *d* (3.2)	4.96, *d* (3.0)
4	5.70, *d* (3.2)	5.67, *d* (3.3)	5.74, *d* (3.2)
5	4.60, *t* (8.4)	4.57, *m*	4.53, *t* (7.8)
6a	4.85, *dd* (8.9, 11.3)	4.61, *m*	4.66, *dd* (7.8, 11.3)
6b	4.27, *dd* (7.9, 11.3)	4.30, *dd* (6.8, 10.4)	4.31, *dd* (7.8, 11.3)
1-*O*-Galloyl			
2′, 6′	7.07, *s*	7.12, *s*	7.15, *s*
4-*O*-Neochebuloyl			
2′	5.50, br *s*	5.56, *d* (1.1)	5.38, br *s*
3′	3.99, *d* (8.3)	3.97, *dd* (0.9, 9.4)	4.20, *d* (4.2)
4′	3.28, *m*	3.28, *td* (9.0, 5.3)	3.55, td (7.4, 3.2)
5′a	2.86, *dd* (10.0, 17.4)	2.80, *dd* (8.6, 17.3)	2.91, *dd* (10.6, 17.1)
5′b	2.46, *dd* (4.6, 17.4)	2.52, *dd* (5.2, 17.3)	2.43, *dd* (3.9, 17.1)
2″	7.05, *s*	7.07, *s*	7.15, *s*
		1′-*O*-Me	1′-*O*-Me
		3.55, *s*	3.62, *s*
		6′-*O*-Me	6′-*O*-Me
		3.52, *s*	3.58, *s*
HHDP			
5′	6.75, *s*	6.85, *s*	6.89, *s*
5″	6.74, *s*	6.71, *s*	6.74, *s*

^{1}H NMR data were measured at [a]500 MHz in CD_3OD and [b]600 MHz in acetone-d_6, respectively.

Table 8 ^{13}C NMR data for compounds **22-24**

	22	**23**	**24**
β-D-Glucose			
1	95.4	94.7	94.7
2	69.6	69.9	70.4
3	71.8	71.8	71.5
4	65.2	65.2	65.3
5	74.3	74.2	74.4
6	65.1	64.5	64.6
1-*O*-Galloyl			
1′	120.7	120.2	120.7

2′, 6′	110.9	110.3	110.5
3′, 5′	146.6	145.9	145.9
4′	140.6	139.5	139.5
7′	166.7	165.7	165.2
4-*O*-Neochebuloyl			
1′	172.6	170.6	170.9
2′	79.1	78.1	77.0
3′	37.0	35.8	35.9
4′	45.3	43.8	44.6
5′	34.8	34.1	32.1
6′	175.5	172.6	172.8
7′	173.6	172.5	171.5
1″	116.4	116.1	117.1
2″	109.3	108.9	108.6
3″	146.9	146.2	146.1
4″	140.7	139.3	139.3
5″	144.2	143.3	142.9
6″	117.6	116.4	116.8
7″	166.8	164.3	163.4
		1′-*O*-Me	1′-*O*-Me
		53.0	53.2
		6′-*O*-Me	6′-*O*-Me
		52.1	52.1
HHDP			
1′	117.4	116.5	116.3
2′	145.8	144.8	144.9
3′	138.4	137.0	137.1
4′	145.4	144.8	144.8
5′	110.2	109.7	109.8
6′	125.3	125.0	125.4
7′	168.3	167.0	166.7
1″	116.5	115.4	115.3
2″	145.3	144.7	144.9
3″	137.9	136.4	136.5
4″	146.2	145.2	145.2
5″	109.0	108.1	108.5
6″	125.5	125.2	125.5
7″	170.2	168.7	168.4

^{13}C NMR data were measured at [a]125 MHz in CD_3OD and [b]150 MHz in acetone-d_6, respectively.

Table 9. ^{1}H and ^{13}C NMR data for compounds **29**-**30**

Position	**29**		**30**	
	δ_H	δ_C	δ_H	δ_C
(-)-shikimic acid				
1		130.5		132.4
2a	2.68, *m*	32.0	2.74, *m*	31.2
2b	2.32, *m*		2.27, *m*	
3	4.27, *dt* (9.7, 4.8)	65.3	4.00, *dd* (3.8, 7.2)	67.8
4	4.69, *t* (3.7)	74.8	4.15, *dt* (7.1, 5.0)	70.1
5	5.11, *dd* (3.8, 7.3)	65.6	5.72, *m*	70.4
6	6.84, *m*	138.8	6.75, *dd* (1.1, 2.3)	134.2
7		169.1		168.3
	4-*O*-Galloyl		5-*O*-Galloyl	
1′		121.7		121.0
2′, 6′	7.08, *s*	110.3	7.14, *s*	109.8
3′, 5′		146.2		145.8
4′		139.2		138.9
7′		167.3		166.7

^{1}H NMR data were measured at 300 MHz in acetone-d_6:D_2O = 9:1. ^{13}C NMR data were measured at 75 MHz in acetone-d_6:D_2O = 9:1.

Table 10. ^{1}H NMR data for compounds **31-33**

	31[a]		**32**[b]		**33**[c]	
Glucose	*α*	*β*	*α*	*β*	*α*	*β*
1	5.10, d (3.7)	4.54, *d* (7.8)	5.20, *d* (3.7)	4.71, *d* (8.0)	5.33, *d* (4.2)	4.87, *d* (7.7)
2	3.41 *m*	3.37, *m*	3.64, *m*	3.47, *dd* (8.0, 9.5)	3.81, *dd* (4.2, 9.8)	3.62, *dd* (7.7, 9.6)
3	3.45, *m*	3.42, *m*	5.39, *t* (9.5)	5.16, *t* (9.0)	5.70, *t* (9.7)	5.51, *t* (9.6)
4	3.46, *m*	3.43, *m*	3.76, m	3.71, *m*	5.34, *t* (9.9)	5.33, *m*
5	4.03, *ddd* (2.1, 4.9, 10.0)	3.58, *m*	4.20, *ddd* (2.1, 5.0, 10.2)	3.79, *m*	4.51, *ddd* (2.1, 5.0, 10.2)	4.12, *ddd* (2.1, 5.3, 10.0)
6	4.49, *m*	4.45, *m*	4.52, *m*	4.51, *m*	4.41, *m*	4.43, *m*
	4.32, *m*	4.27, *m*	4.41, *m*	4.38, *m*	4.27, *t* (5.0)	4.25, *t* (5.4)
6-*O*-Gal						
2′, 6′	7.10, *s*	7.10, *s*	7.12, *s*	7.12, *s*	7.02, *s*	7.02, *s*
3-*O*-Gal						

2″, 6″		7.13, *s*	7.13, *s*	7.14, *s*	7.13, *s*
4-*O*-Gal					
2‴, 6‴				7.01, *s*	6.99, *s*

[1]H NMR data were measured at [a]300, [b]400, and [c]600 MHz in acetone-d_6:D_2O = 9:1, respectively.

Table 11. [1]H NMR data for compounds **34**-**37**

Position	**34**[a]	**35**[b]	**36**[b]	**37**[a]
2	5.08, *s*	5.25, *d* (0.9)	5.35, *d* (0.9)	4.92, *s*
3	3.56, *d* (8.4)	3.90, *dd* (0.9, 9.1)	3.84, *dd* (0.9, 8.7)	3.94, *d* (4.4)
4	2.82, *m*	3.17, *td* (9.1, 5.5)	3.12, *td* (4.6, 9.8)	3.20, m
5a	2.67, *dd* (11.3, 16.9)	2.84, *dd* (9.1, 16.9)	2.85, *dd* (10.2, 17.1)	2.57, *dd* (11.9, 16.8)
5b	2.06, *dd* (3.5, 16.9)	2.46, *dd* (5.5, 16.9)	2.35, *dd* (4.6, 17.1)	1.87, *dd* (2.8, 16.8)
2′	6.87, *s*	7.03, *s*	7.04, *s*	6.96, *s*
		OMe	OMe	
		3.52, *s*	3.62, *s*	

[1]H NMR data were measured at 400 MHz in [a]DMSO-d_6 and [b]CD_3OD, respectively.

Table 12. [13]C NMR data for compounds **34**-**37**

Position	**34**[a]	**35**[b]	**36**[b]	**37**[a]
1	170.8	172.5	171.3	171.3
2	76.3	78.7	78.7	76.2
3	35.6	37.1	37.1	34.9
4	43.7	45.2	45.2	43.4
5	34.1	35.1	35.1	31.5
6	172.9	173.9	175.4	173.1
7	174.4	176.6	176.7	174.2
1′	114.8	116.4	116.0	115.4
2′	107.6	109.2	109.2	107.2
3′	145.3	146.7	146.8	145.5
4′	138.7	140.6	140.7	139.2
5′	142.7	144.1	144.0	142.4

6′	116.5	117.7	117.6	116.3
7′	163.3	166.7	166.3	163.4
		OMe	OMe	
		52.1	53.3	

^{13}C NMR data were measured at 100 MHz in [a]DMSO-d_6 and [b]CD_3OD, respectively.

Table 13. ^{1}H and ^{13}C NMR data for compounds **42**-**43**

Position	**42**[a]		**43**[b]	
	δ_H	δ_C	δ_H	δ_C
β-D-Glucose				
1	6.36, *d* (1.8)	94.9	6.21, *d* (4.6)	95.0
2	3.98, *d* (1.8)	69.3	4.22, br *d* (4.6)	70.8
3	4.80, br *s*	71.4	4.88, br *s*	73.2
4	4.46, *d* (3.0)	62.3	5.67, *d* (3.1)	65.2
5	4.16, *dd* (2.4, 8.4)	76.0	4.50, *m*	75.4
6a	4.95, *t* (10.2)	64.9	4.51, *m*	65.0
6b	4.52, *t* (10.2)		4.30, *dd* (10.8, 14.0)	
1-*O*-Galloyl				
1′		120.4		121.0
2′, 6′	7.05, *s*	110.8	7.12, *s*	110.6
3′, 5′		146.2		146.4
4′		140.2		139.8
7′		166.5		166.4
HHDP				
1′		117.0		117.0
2′		145.4		145.3
3′		138.0		137.4
4′		145.1		145.2
5′	6.68, *s*	110.1	6.79, *s*	110.1
6′		125.3		125.5
7′		168.4		167.6
1″		116.5		116.1
2″		145.0		145.3
3″		137.5		137.0
4″		145.9		145.7
5″	6.66, *s*	108.2	6.71, *s*	108.8
6″		125.4		125.6
7″		170.0		169.1
			4-*O*-Galloyl	
1′				120.7
2″, 6″			7.08, *s*	110.8
3″, 5″				146.3

4″	140.1
7″	166.3

[1]H NMR data were measured at 400 MHz in [a]CD_3OD and [b]acetone-d_6, respectively.

[13]C NMR data were measured at 100 MHz in [a]CD_3OD and [b]acetone-d_6, respectively.

Table 14. [1]H NMR data for compounds **44**-**45**

	44[a]		**45**[b]	
Glucose	*α*	*β*	*α*	*β*
1	5.25, *d* (3.8)	4.72, *d* (7.8)	5.56, *d* (3.7)	5.08, *d* (8.1)
2	3.81, *m*	3.58, *dd* (7.8, 9.6)	5.12, *m*	5.24, *dd* (6.4, 8.1)
3	5.46, *t* (9.2)	5.28, *t* (9.6)	5.88, *t* (10.0)	5.61, *t* (9.8)
4	4.92, *t* (10.4)	4.95, *t* (10.4)	5.11, *m*	5.11, *m*
5	4.53, *dd* (6.1, 10.2)	4.07, *dd* (6.1, 9.8)	4.67, *dd* (6.4, 9.8)	4.27, *dd* (6.0, 9.9)
6a	5.18, *m*	5.20, *m*	5.29, *m*	5.29, m
6b	3.74, *m*	3.81, *m*	3.77, *dd* (1.0, 13.0)	3.85, br *d* (13.1)
HHDP				
5′	6.46, *s*	6.45, *s*	6.47, *s*	6.44, *s*
5″	6.60, *s*	6.61, *s*	6.65, *s*	6.66, *s*
		3-*O*-Galloyl		2-*O*-Galloyl
2′, 6′	7.02, *s*	7.01, *s*	7.06, *s*	7.05, *s*
				3-*O*-Galloyl
2″, 6″			6.98, *s*	6.94, *s*

[1]H NMR data were measured at 400 MHz in [a]CD_3OD and [b] acetone-d_6:D_2O = 9:1, respectively.

Table 15. [1]H NMR data for compounds **48**-**49**

	48[a]		**49**[b]	
Glucose	*α*	*β*	*α*	*β*
1	5.11, *d* (3.5)	4.72, *d* (8.1)	5.30, *m*	4.92, *m*
2	4.80, *m*	4.64, *dd* (8.1, 9.5)	4.85, *m*	4.66, *dd* (8.2, 9.3)
3	5.22, *t* (9.6)	4.90, *t* (9.0)	5.28, *m*	5.02, *t* (9.7)
4	4.77, *t* (9.6)	4.81, *m*	5.27, *m*	5.27, *m*
5	3.27, *td*	2.70, *td*	4.13, *m*	3.80, *m*

	(10.3, 1.4)	(10.1, 1.6)		
6a	4.21, *t* (7.0)	4.18, *t* (7.2)	4.41, *dd* (1.7, 12.0)	4.51, *dd* (1.6, 12.2)
6b	2.08, m	2.16, *dd* (1.4, 11.2)	4.04, *m*	4.07, *m*
HHDP				
5′	6.52, *s*	6.52, *s*	6.39, *s*	6.37, *s*
5″	6.58, *s*	6.65, *s*	6.47, *s*	6.48, *s*
		Gallagyl		
2′	6.58, *s*	6.59, *s*		
2″	7.02, *s*	7.02, *s*		
			Flavogallonyl	
2′			7.20, *s*	7.16, *s*
5‴			7.44, *s*	7.43, *s*
			6-*O*-Galloyl	
2′, 6′			6.84, *s*	6.82, *s*

[1]H NMR data were measured at [a]600 and [b]400 MHz in acetone-d_6:D_2O = 9:1, respectively.

Table 16. [1]H NMR data for compounds **50-52**

Position	**50**[a]	**51**[b]	**52**[b]
1a	2.31, *dd* (4.1, 12.3)	2.19, *dd* (4.2, 11.2)	2.19, *m*
1b	1.43, *m*	1.42, *t* (11.9)	1.52, *m*
2	4.28, *td* (10.5, 4.1)	4.20, *m*	4.21, *m*
3	4.22, *d* (9.1)	4.11, *d* (9.1)	4.08, *d* (9.1)
5	2.11, *m*	1.90, *m*	2.00, *m*
6a	1.76, *m*	1.90, *m*	1.82, br *d* (11.7)
6b	1.50, *td* (3.4, 12.3)	1.52, *m*	1.53, *m*
7a	1.70, *m*	2.12, *t* (11.4)	2.13, *m*
7b	1.35, *m*	1.64, *d* (12.6)	1.50, *m*
9	1.88, *d* (11.5)	1.90, *m*	2.29, *m*
11	2.14, *m*	2.05, *m*	2.05, *m*
12	5.57, br *t* (4.1)	5.47, br *t* (4.2)	5.54, br *t* (4.0)
15a	2.18, *m*	2.34, *td* (13.2, 3.2)	2.08, *m*
15b	1.17, *m*	1.35, *d* (13.6)	1.30, *d* (12.8)
16a	2.82, *td* (3.6, 12.70)	2.70, *td* (12.8, 2.4)	2.77, br *t* (12.1)
16b	2.10, *m*	1.97, *m*	2.03, *m*
18	3.61, br *s*	3.52, br *s*	3.61, br *s*
19	3.63, br *s*	3.55, br *s*	3.59, br *s*
21a	2.12, *m*	1.99, *m*	2.10, *m*
21b	1.15, *m*	1.02, *m*	1.16, *m*
22a	2.20, *m*	2.06, *m*	2.20, *m*
22b	2.04, *m*	1.93, *m*	2.03, *m*

23a	4.22, *d* (10.5)	4.92, *d* (11.1)	4.97, *d* (11.0)
23b	3.74, *d* (10.5)	4.41, *d* (11.1)	4.32, *d* (11.0)
24	1.12, *s*	1.06, *s*	1.06, *s*
25	1.09, *s*	1.08, *s*	1.07, *s*
26	1.11, *s*	1.18, *s*	1.10, *s*
27	1.59, *s*	1.89, *s*	1.88, *s*
29	1.20, *s*	1.14, *s*	1.20, *s*
30	1.12, *s*	1.00, *s*	1.13, *s*
		Neochebuloyl	4'-*epi*-Neochebuloyl
2'		6.00, *s*	5.80, *s*
3'		5.03, *d* (8.3)	5.32, *d* (4.0)
4'		4.12, *m*	4.52, br *d* (10.6)
5'a		3.75, *dd* (10.3, 17.0)	3.70, *dd* (11.2, 16.8)
5'b		3.23, *dd* (3.7, 17.0)	3.09, *d* (16.8)
2''		7.89, *s*	7.94, *s*
		β-D-Glucose	
1'''		6.39, *d* (8.1)	
2'''		4.21, *t* (8.4)	
3'''		4.31, *t* (8.9)	
4'''		4.40, *m*	
5'''		4.05, *m*	
6'''a		4.46, *m*	
6'''b		4.43, *m*	

[1]H NMR data were measured at [a]600 and [b]800 MHz in pyridine-d_5, respectively.

Table 17. [13]C NMR data for compounds **50**-**52**

Position	**50**	**51**	**52**
1	48.0	47.6	47.4
2	69.3	69.1	69.2
3	78.7	78.0	77.7
4	44.2	43.6	43.6
5	49.0	49.3	49.1
6	19.2	19.5	19.5
7	33.5	33.4	33.5
8	40.6	40.8	40.6
9	48.5	48.7	48.7
10	39.1	39.0	39.0
11	24.8	24.7	24.8
12	123.5	123.9	124.2
13	145.4	145.2	145.6
14	42.7	42.8	42.9
15	29.6	29.2	29.3

16	28.8	28.5	29.0
17	46.5	47.0	46.6
18	45.3	45.1	45.4
19	81.6	81.6	81.9
20	36.2	36.1	36.2
21	29.6	29.5	29.8
22	34.2	33.5	34.1
23	66.9	68.6	68.4
24	14.8	14.9	14.4
25	17.8	17.8	17.7
26	18.1	18.2	18.1
27	25.3	25.6	25.3
28	181.4	177.8	181.4
29	29.4	29.2	29.2
30	25.3	25.2	25.5
		Neochebuloyl	4'-*epi*-Neochebuloyl
1'		173.5	174.1
2'		79.5	78.9
3'		37.9	37.3
4'		46.6	45.9
5'		36.1	33.7
6'		175.0	175.3
7'		174.9	174.6
1''		117.2	118.5
2''		109.4	109.0
3''		147.8	147.8
4''		141.8	142.0
5''		145.6	145.3
6''		118.7	118.5
7''		165.6	165.6
		β-D-Glucose	
1'''		96.3	
2'''		74.7	
3'''		79.5	
4'''		71.6	
5'''		79.8	
6'''		62.6	

^{13}C NMR data were measured at [a]150 and [b]200 MHz in pyridine-d_5, respectively.

Table 18. ^{1}H and ^{13}C NMR data for compounds **53-54**

Position	**53**[a]		**54**[b]	
	δ_H	δ_C	δ_H	δ_C

1a	2.23, *m*	48.0	2.23, *dd* (4.3, 12.6)	48.1
1b	1.24, *m*		1.21, *m*	
2	4.22, *m*	68.9	4.22, *m*	68.9
3	3.94, *d* (9.3)	78.2	3.91, *d* (9.3)	78.3
4		43.8		43.8
5	1.80, *m*	49.7	1.78, br *d* (12.1)	50.0
6a	1.68, *m*	19.4	1.70, *m*	19.5
6b	1.44, *m*		1.47, *m*	
7a	1.70, *m*	33.6	1.73, *m*	33.5
7b	1.32, *m*		1.41, *m*	
8		40.5		40.7
9	2.06, m	49.3	2.05, *m*	49.3
10		38.8		38.9
11	2.05, m	24.7	2.04, *m*	24.7
12	5.55, br *t* (4.0)	123.5	5.50, br *t* (4.0)	124.3
13		145.5		144.9
14		42.6		42.6
15a	2.16, *m*	29.3	2.29, *td* (13.9, 3.0)	29.3
15b	1.18, *m*		1.17, *m*	
16a	2.77, *td* (14.0, 3.2)	28.9	2.76, *td* (13.2, 3.4)	28.5
16b	2.08, *m*		2.08, *m*	
17		46.6		47.0
18	3.63, br *s*	45.3	3.52, br *s*	45.1
19	3.61, br *s*	81.8	3.57, br *s*	81.6
20		36.2		36.0
21a	2.05, *m*	29.7	2.05, *m*	29.5
21b	1.15, *m*		1.06, *m*	
22a	2.17, *m*	34.1	2.06, *m*	33.5
22b	2.02, *m*		1.96, *m*	
23a	4.74, *d* (11.1)	67.6	4.71, *d* (11.1)	68.0
23b	4.51, *d* (11.1)		4.54, *d* (11.1)	
24	1.06, *s*	14.4	1.09, *s*	14.4
25	1.06, *s*	17.6	1.09, *s*	17.7
26	1.19, *s*	18.0	1.16, *s*	18.1
27	1.72, *s*	25.4	1.66, *s*	25.3
28		181.4		177.7
29	1.12, *s*	29.3	1.14, *s*	29.2
30	1.05, *s*	25.2	0.99, *s*	25.2
	Galloyl		Galloyl	
1'		122.0		122.1
2',6'	7.89, s	110.6	7.85, *s*	110.6
3',5'		148.2		148.2
4'		141.5		141.4
7'		167.7		167.6
			β-D-Glucose	
1"			6.38, *d* (8.2)	96.4

2"		4.22, *t* (8.6)	74.6
3"		4.30, *t* (8.9)	79.5
4"		4.38, *t* (9.2)	71.6
5"		4.05, *m*	79.8
6"a		4.47, *dd* (2.4, 11.9)	62.7
6"b		4.42, *dd* (4.3, 11.9)	

^{1}H NMR data were measured at [a]400 and [b]600 MHz in pyridine-d_5, respectively.

^{13}C NMR data were measured at [a]100 and [b]150 MHz in pyridine-d_5, respectively.

Table 19. ^{1}H and ^{13}C NMR data for compounds **56**-**57**

Position	**56**		**57**	
	δ_H	δ_C	δ_H	δ_C
1a	2.39, *dd* (3.8, 12.1)	50.6	2.32, *dd* (3.8, 12.0)	50.6
1b	1.50, *t* (11.8)		1.30, *m*	
2	4.43, *m*	69.6	4.38, *m*	69.0
3	4.25, *d* (9.3)	78.9	4.04, *d* (9.4)	78.6
4		45.0		44.4
5	2.01, *m*	49.4	1.94, *m*	50.2
6	5.11, br *s*	68.1	4.89, br *s*	68.1
7a	2.01, *m*	41.6	1.93, *m*	41.7
7b	1.84, *m*		1.84, *m*	
8		39.9		39.9
9	2.10, *m*	49.3	1.94, *m*	49.6
10		38.7		38.6
11a	2.09, *m*	24.2	2.00, *m*	24.0
11b	1.94, br *d* (12.8)		1.89, *m*	
12	5.58, br *s*	123.4	5.52, br *s*	123.4
13		144.7		144.0
14		43.3		
15a	2.33, *m*	28.8	2.38, br *t* (12.9)	28.5
15b	1.22, *m*		1.11, *m*	
16a	2.29, *m*	24.6	2.24, *m*	24.5
16b	2.20, *m*		2.08, *m*	
17		47.2		47.5
18	3.35, *d* (10.9)	42.6	3.21, *d* (11.4)	42.4
19a	1.80, *m*	46.9	1.72, *m*	46.7
19b	1.32, br *d* (11.8)		1.25, *m*	
20		31.5		31.2
21a	1.45, *m*	34.7	1.31, *m*	34.5
21b	1.20, *td* (13.8, 3.0)		1.07, *m*	
22a	2.05, *m*	33.7	1.81, *m*	32.9
22b	1.81, *m*		1.73, *m*	
23a	4.42, *d* (10.4)	66.7	4.94, *d* (10.9)	67.6

23b	4.06, *d* (10.4)		4.80, *d* (10.9)	
24	1.76, *s*	16.4	1.78, *s*	16.1
25	1.80, *s*	19.5	1.78, *s*	19.5
26	1.64, *s*	19.1	1.70, *s*	19.3
27	1.25, *s*	26.8	1.19, *s*	26.4
28		180.7		176.9
29	0.94, *s*	33.7	0.91, *s*	33.6
30	1.02, *s*	24.3	0.88, *s*	24.1
			Galloyl	
1'				121.9
2',6'			7.84, *s*	110.5
3',5'				148.1
4'				141.5
7'				167.7
			β-D-Glucose	
1"			6.29, *d* (7.9)	96.3
2"			4.18, *t* (8.4)	74.5
3"			4.27, *t* (8.7)	79.2
4"			4.37, *t* (9.6)	71.6
5"			4.02, *m*	79.7
6"a			4.45, *dd* (2.2, 11.6)	62.6
6"b			4.40, *m*	

^{1}H NMR data were measured at 600 MHz in pyridine-d_5. ^{13}C NMR data were measured at 150 MHz in pyridine-d_5.

Table 20. ^{1}H and ^{13}C NMR data for compounds **58**-**59**

Position	**58**		**59**	
	δ_H	δ_C	δ_H	δ_C
1a	2.32, *dd* (3.8, 12.2)	48.2	2.32, *dd* (3.7, 12.6)	48.3
1b	1.40, *m*		1.40, *m*	
2	4.25, *m*	69.4	4.26, *m*	69.4
3	4.22, *d* (8.2)	78.6	4.21, *d* (8.2)	78.7
4		44.2		44.2
5	1.85, *m*	48.4	1.83, *m*	48.5
6a	1.76, *m*	19.0	1.72, *m*	19.1
6b	1.46, *m*		1.45, *m*	
7a	1.66, *m*	33.7	1.65, *m*	33.3
7b	1.33, *m*		1.37, *m*	
8		40.3		40.5
9	1.93, *m*	48.7	1.90, *m*	48.7
10		38.9		38.9

11	2.06, *m*	24.2	2.03, *m*	24.5
12	5.48, br *t* (3.2)	123.0	5.44, br *s* (3.2)	123.4
13		145.3		144.7
14		42.7		42.7
15	2.17, *m*	28.8	2.35, *m*	28.7
	1.16, *m*		1.13, *m*	
16	2.06, *m*	24.5	2.05, *m*	23.9
	1.96, *m*		1.94, br *d* (13.0)	
17		47.1		47.5
18	3.30, *dd* (4.1, 13.8)	42.4	3.20, *dd* (3.7, 13.7)	42.3
19	1.78, *m*	46.8	1.72, *m*	46.7
	1.27, *m*		1.23, *m*	
20		31.4		31.3
21	1.36, *m*	34.7	1.33, *m*	34.5
	1.13, *m*		1.08, *m*	
22	1.83, *m*	33.4	1.83, *m*	33.1
	1.67, *m*		1.74, *m*	
23	4.24, *d* (10.5)	66.9	4.22, *d* (10.5)	67.0
	3.74, *d* (10.5)		3.73, *d* (10.5)	
24	1.09, *s*	14.9	1.08, *s*	14.9
25	1.06, *s*	18.0	1.11, *s*	18.0
26	1.09, *s*	17.9	1.17, *s*	18.1
27	1.22, *s*	26.6	1.17, *s*	26.6
28		180.7		177.0
29	0.93, *s*	33.7	0.89, *s*	33.6
30	1.11, *s*	24.1	0.89, *s*	24.2
			β-D-Glucose	
1'			6.34, *d* (8.1)	96.3
2'			4.22, *t* (8.6)	74.6
3'			4.31, *t* (8.9)	79.4
4'			4.38, *t* (9.2)	71.7
5'			4.05, *m*	79.9
6'a			4.49, br *d* (10.5)	62.8
6'b			4.42, *dd* (4.3, 10.9)	

^{1}H NMR data were measured at 400 MHz in pyridine-d_5. ^{13}C NMR data were measured at 100 MHz in pyridine-d_5.

Table 21. ^{1}H and ^{13}C NMR data for compounds **60-61**

Position	**60**		**61**	
	δ_H	δ_C	δ_H	δ_C

1a	2.19, *dd* (4.4, 12.7)	48.2	2.24, *dd* (4.1, 12.4)	48.3
1b			1.21, *m*	
2	4.17, *td* (10.3, 4.3)	68.8	4.21, *m*	68.8
3	3.95, *d* (9.3)	78.2	3.97, *d* (9.4)	78.2
4		43.7		43.7
5	1.70, *m*	49.2	1.73, *m*	49.5
6a	1.59, *m*	19.2	1.63, *m*	19.2
6b	1.36, *m*		1.41, *m*	
7a	1.57, *m*	33.5	1.61, *m*	33.4
7b	1.21, *m*		1.31, *m*	
8		40.2		40.4
9	1.74, *m*	49.0	1.77, *m*	49.1
10		38.7		38.7
11	1.91, *m*	24.4	1.96, *m*	24.4
12	5.41, br *t* (3.3)	122.7	5.41, br *t* (3.3)	122.0
13		145.4		144.7
14		42.6		42.6
15	2.02, *m*	28.5	2.27, *m*	28.5
	1.00, *m*		1.07, *m*	
16	1.99, *m*	24.1	1.97, *m*	23.9
	1.85, *m*		1.90, br *d* (13.3)	
17		47.1		47.5
18	3.23, *dd* (4.6, 14.1)	42.5	3.18, *dd* (4.0, 13.7)	42.3
19	1.67, *m*	46.8	1.69, *m*	46.6
	1.19, *m*		1.21, *m*	
20		31.4		31.2
21	1.36, *m*	34.7	1.32, *m*	34.5
	1.14, *m*		1.08, *m*	
22	1.97, *m*	33.6	1.82, *m*	33.0
	1.75, *m*		1.73, *m*	
23	4.67, *d* (11.0)	67.2	4.71, *d* (11.1)	67.5
	4.50, *d* (11.0)		4.53, *d* (11.1)	
24	1.00, s	14.5	1.06, *s*	14.1
25	0.94, *s*	17.9	1.05, *s*	17.8
26	0.94, *s*	17.7	1.08, *s*	18.1
27	1.13, *s*	26.4	1.15, *s*	26.3
28		180.6		176.9
29	0.89, *s*	33.7	0.90, *s*	33.8
30	0.94, *s*	24.2	0.88, *s*	24.1
	Galloyl		Galloyl	
1'		122.0		122.0
2',6'	7.85, *s*	110.6	7.87, *s*	110.5
3',5'		148.2		148.1
4'		141.5		141.4
7'		167.6		167.6
			β-D-Glucose	

1"	6.33, *d* (8.1)	96.2
2"	4.20, *t* (8.7)	74.6
3"	4.29, *t* (8.9)	79.4
4"	4.37, *t* (9.2)	71.6
5"	4.05, *m*	79.8
6"a	4.48, *dd* (2.2, 11.9)	62.7
6"b	4.42, *dd* (4.4, 11.9)	

^{1}H NMR data were measured at 600 MHz in pyridine-d_5. ^{13}C NMR data were measured at 150 MHz in pyridine-d_5.

Table 22. ^{1}H and ^{13}C NMR data for compounds **62-63**

Position	**62**		**63**	
	δ_H	δ_C	δ_H	δ_C
1a	2.23, *m*	48.0	2.27, *dd* (4.3, 12.4)	48.1
1b	1.27, *m*		1.30, *m*	
2	4.08, *td* (10.3, 4.3)	69.1	4.13, *td* (11.0, 4.3)	69.1
3	3.36, *d* (9.4)	84.3	3.40, *d* (9.4)	84.3
4		40.4		40.3
5	1.03, *m*	56.5	1.06, *m*	56.5
6a	1.56, *m*	19.5	1.58, *m*	19.6
6b	1.42, *m*		1.47, *m*	
7a	1.54, *m*	33.8	1.59, *m*	33.7
7b	1.35, *m*		1.49, *m*	
8		40.6		40.8
9	1.97, *m*	48.9	1.99, *m*	48.9
10		39.2		39.2
11	2.08, *m*	24.8	2.10, *m*	24.8
12	5.53, br *t* (3.2)	123.9	5.52, br *t* (3.2)	125.0
13		145.4		144.9
14		42.7		42.6
15a	2.17, *m*	29.8	2.41, *td* (13.8, 3.0)	29.5
15b	1.22, *m*		1.30, *m*	
16a	2.82, *m*	28.8	2.86, *td* (13.3, 3.4)	28.5
16b	2.12, *m*		2.16, *m*	
17		46.5		47.0
18	3.58, br *s*	45.3	3.55, br *s*	45.1
19	3.60, br *s*	81.7	3.59, br *s*	81.5
20		36.2		36.0
21a	2.10, *m*	29.6	2.07, *m*	29.4
21b	1.13, *m*		1.06, *m*	
22a	2.17, *m*	34.1	2.08, *m*	33.5

22b	2.02, *m*		1.98, *m*	
23	1.27, *s*	29.7	1.28, *s*	29.8
24	1.10, *s*	18.0	1.11, *s*	18.1
25	1.06, *s*	18.1	1.07, *s*	18.1
26	1.01, *s*	17.3	1.19, *s*	17.4
27	1.61, *s*	25.3	1.63, *s*	25.4
28		181.4		177.7
29	1.17, *s*	29.3	1.16, *s*	29.2
30	1.10, *s*	25.3	1.00, *s*	25.1
			β-D-Glucose	
1'			6.41, *d* (8.1)	95.7
2'			4.25, *t* (9.4)	73.9
3'			4.32, *t* (8.9)	78.7
4'			4.40, *m*	70.9
5"			4.06, *m*	79.1
6'			4.46, *m*	61.9

^{1}H NMR data were measured at 600 MHz in pyridine-d_5. ^{13}C NMR data were measured at 150 MHz in pyridine-d_5.

Table 23. ^{1}H and ^{13}C NMR data for compounds **65-66**

Position	**65**[a]		**66**[b]	
	δ_H	δ_C	δ_H	δ_C
1a	2.41, *dd* (4.2, 12.4)	48.9	2.29, *dd* (4.1, 12.6)	48.8
1b	1.43, *m*		1.15, *m*	
2	4.28, *m*	69.5	4.22, *m*	68.9
3	4.23, *d* (9.7)	78.7	3.95, *d* (9.3)	78.1
4		44.1		43.7
5	1.86, *d* (11.8)	48.6	1.74, *d* (11.9)	49.4
6a	1.72, *m*	18.9	1.60, *m*	19.0
6b	1.43, *m*		1.37, *m*	
7a	1.75, *m*	35.2	1.70, *m*	35.2
7b	1.46, *m*		1.39, *m*	
8		40.1		40.0
9	1.74, *m*	48.9	1.63, *m*	49.1
10		38.7		38.5
11	2.10, *m*	24.2	2.02, *m*	24.1
12	5.71, br *t* (3.8)	127.1	5.68, *dd* (2.4, 4.8)	126.9
13		139.3		139.4
14		45.5		45.4
15a	2.45, *m*	29.5	2.40, *td* (13.6, 3.4)	29.3
15b	1.18, *m*		1.13, *m*	

16a	2.19, *td* (13.8, 3.0)	31.5	2.17, br *d* (13.7)	31.4
16b	1.68, *td* (14.4, 3.0)		1.67, *m*	
17		50.4		50.4
18		134.3		134.3
19		136.4		136.4
20	2.13, *m*	35.1	2.14, *dt* (15.1, 7.7)	35.1
21a	2.09, *m*	27.3	2.06, *m*	27.2
21b	1.26, br *d* (12.8)		1.25, *m*	
22a	2.56, *td* (12.6, 3.0)	35.8	2.52, *td* (12.8, 3.2)	35.7
22b	1.49, *m*		1.43, *m*	
23a	4.23, *d* (10.4)	67.0	4.75, *d* (11.1)	67.3
23b	3.74, *d* (10.4)		4.50, *d* (11.1)	
24	1.08, *s*	14.9	1.04, *s*	14.7
25	1.14, *s*	18.7	1.05, *s*	18.5
26	1.22, *s*	19.1	1.15, *s*	19.1
27	1.03, *s*	22.6	1.09, *s*	22.4
28		175.2		175.2
29	1.79, *s*	20.0	1.77, *s*	20.1
30	1.02, *d* (7.0)	19.1	1.03, *d* (7.0)	19.2
	β-D-Glucose		*β*-D-Glucose	
1'	6.35, *d* (8.2)	96.3	6.34, *d* (8.2)	96.3
2'	4.17, *t* (8.6)	74.6	4.16, *t* (8.6)	74.6
3'	4.28, *t* (9.0)	79.4	4.28, *t* (9.0)	79.4
4'	4.36, *t* (9.1)	71.7	4.35, *t* (9.3)	71.6
5'	4.00, *m*	79.6	4.00, *m*	79.7
6'a	4.40, *m*	62.8	4.41, *dd* (2.0, 11.6)	62.8
6'b			4.38, *dd* (4.2, 11.6)	
			Galloyl	
1"				121.9
2",6"			7.88, *s*	110.5
3",5"				148.2
4"				141.4
7"				167.6

^{1}H NMR data were measured at [a]600 and [b]800 MHz in pyridine-d_5, respectively. ^{13}C NMR data were measured at [a]150 and [b]200 MHz in pyridine-d_5, respectively.

2.5. General partial hydrolysis and acid hydrolysis

Partial hydrolysis: Solutions of compound **3**, **18**, **23**, and **24** (each 4 mg) in water were heated to 80℃ for 10 h. The reaction mixtures were submitted to HPLC analysis to reveal the presence of corilagin (**42**, t_R 24.2 min), and corilagin (**42**) was

purified from each reaction mixtures by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 11:89).

Acid hydrolysis: Compounds **3**, **20**, **21**, **23**, **24** (each 4.0 mg), **51**, and **52** (each 0.5 mg) were dissolved in 1N H_2SO_4 (2 ml) and solution was heated to 80℃ for 4 h. After neutralizing acidic solution with saturated Na_2CO_3 solution, the solvent was evaporated to dryness under nitrogen gas. The reaction mixtures were submitted to HPLC analysis to reveal the presence of chebulic acid (**34**, t_R 5.7 min, compound **3**, **20**, **21**, **23**, and **51**) or neochebulic acid (**37**, t_R 5.2 min, compound **24** and **52**), respectively. The chebulic acid and neochebulic acid were purified from each reaction mixtures by semi-preparative HPLC (MeCN/H_2O with 0.1% formic acid 1:99).

2.6. Determination of absolute configuration of sugar

Compounds **3**, **11**, **12**, **13**, **18**, **20**, **21**, **23**, **24** (each 1 mg), **51**, and **66** (each 0.5 mg) were hydrolyzed by 1 N H_2SO_4 (1 ml) heated at 80℃ for 4 h and neutralized with saturated aqueous Na_2CO_3 solution. After the solution was dried *in vacuo*, each residue was dissolved in pyridine (0.2 ml) containing L-cysteine methyl ester hydrochloride (1.0 mg) and heated at 60℃ for 1 h. One microliter of *o*-torylisothiocyanate was added to the mixtures, which were heated at 60℃ for 1 h. Each final mixture was directly analyzed by analytical RP-HPLC (solvent system: aqueous 0.1% formic acid (A) and MeCN (B), gradient condition: 20% to 40% B in 40 min, 1 ml/min). The t_R of the peak at 20.9 and 27.6 min coincided with that of the thiocarbamoyl thiazolidine derivatives of D-glucose and L-rhamnose, respectively.

2.7. Evaluation of Baker's yeast α-glucosidase inhibitory activity

The assay to test the inhibition of the Baker's yeast α-glucosidase activity was performed as described by Escandon-Rivera et al. with some modifications (Escandon-Rivera et al., 2012). Briefly, 80 μL of 0.11 U/mL α-glucosidase (from *Saccharomyces cerevisiae*) in 50 mM PBS buffer (pH 7.0) and 40 μL of the samples were mixed and incubated at 36˚C for 10 min in a 96-well microtiter plate. Then, 80 μL of 1.35 mM *p*-nitrophenyl-*α*-D-glucopyranoside (PNPG) was added to initiate the reaction. The reaction mixture was incubated at 36˚C for 10 min and monitored using a microplate reader at a wavelength of 405 nm. The absorbance of the mixture was measured every 2.5 min up to 5 min. The reaction mixture with the test solution replaced by equivalent PBS buffer was used as a control, and acarbose was used as a positive control. Inhibition of α-glucosidase was calculated with the following equation: % inhibition = $[1-(\Delta Abs_{sample}/\Delta Abs_{control})]\times 100$. All the samples were measured in triplicate. The half-maximal inhibitory concentration (IC_{50}) value was calculated using GraphPad Prism 5 (Graphpad Software, CA, USA) with a dose-inhibition curve with at least six appropriate concentrations. The mode of inhibition of samples was measured with various concentrations of PNPG as a substrate in the absence or presence of acarbose or compound. The inhibition mode was determined by Lineweaver-Burk plot analysis. The initial velocity was the rate of absorbance at 405 nm per 1 min in the assay. Inhibition constants (Ki) were determined by Dixon plot. All the assays were conducted in triplicate.

2.8. Evaluation of rat intestinal α-glucosidase inhibitory activity

The inhibitory activity of all isolated compounds against rat intestinal α-glucosidase was determined by the method of Shai et al., with some modifications (Shai et al., 2011). Twenty microliter of the sample solution was added in a 96-well

microtiter plate. Then, 40 μL of 0.90 mg/mL of PNGP was added into each well. Rat intestinal acetone powder (2 mg/mL in buffer, pH 6.9) was sonicated for 30 min. The suspension was centrifuged at 10,000 g for 15 min. The supernatant was used as α-glucosidase enzyme solution. The reaction was initiated by addition of 40 μL of the enzyme to the reaction mixture in microtiter plates. The plates were incubated for 30 min at 36℃ and then monitored using a microplate reader at a wavelength of 405 nm. The reaction mixture with the sample solution replaced by equivalent PBS buffer was used as a control, and acarbose was used as a positive control. Inhibition of α-glucosidase was calculated with the following equation: % inhibition = $[1-(\Delta Abs_{sample}/\Delta Abs_{control})]\times 100$. All the samples were measured in triplicate.

2.9. Evaluation of porcine pancreatic α-amylase inhibitory activity

The assay to test the inhibition of the α-amylase activity was performed as described by Kusano et al. with some modifications (Kusano et al., 2011). The substrate solution was prepared as follows: potato starch (250 mg) was suspended in 50 mL of 50 mM PBS buffer (pH 7.0) and heated for 20 min at 90˚C. The test sample (20 μL) and substrate (40 μL) solutions were mixed in a 96-well microtiter plate, and the mixtures were incubated at 36˚C for 3 min. Then, 20 μL of porcine pancreas amylase solution (60 U/mL) were added to each well, and the plate was further incubated for 15 min. Forty microliters of 0.2 M HCl was added to terminate the reaction, followed by the addition of 100 μL of 2 mM iodine solution. The absorbances were measured at 650 nm. The inhibitory activity (%) was calculated as follows:

(%) Inhibition = [{1-(Abs 1-Abs 2)/(Abs 3- Abs 4) × 100]

where Abs 1 is the absorbance of the incubated solution containing the sample and substrate; Abs 2 is the absorbance of the incubated solution containing the sample, substrate, and enzyme; Abs 3 is the absorbance of the incubated solution containing the substrate, and Abs 4 is the absorbance of the incubated solution containing the substrate and enzyme.

2.10. Evaluation of protein-tyrosine phosphatase 1B (PTP-1B) inhibitory activity

The inhibitory activity of all isolated compounds against PTP-1B was determined by the method of Abdjul et al., with some modifications (Abdjul et al., 2016). Briefly, 40 μL of 0.25 μg/mL of PTP-1B in 50 mM citrate buffer (pH 6.0) containing 1 mM dithiothreitol, 1 mM EDTA, and 0.1 M NaCl and 20 μL of the samples were mixed and incubated at 36˚C for 10 min in a 96-well microtiter plate. Then, 40 μL of 10 mM *p*-nitrophenylphosphate (pNPP) was added to initiate the reaction. The reaction mixture was incubated at 36˚C for 20 min and monitored using a microplate reader at a wavelength of 405 nm. The absorbance of the mixture was measured every 10 min up to 20 min. The reaction mixture with the test solution replaced by equivalent citrate buffer was used as a control, and oleanolic acid was used as a positive control. Inhibition of PTP-1B was calculated with the following equation: % inhibition = $[1-(\Delta Abs_{sample}/\Delta Abs_{control})]\times 100$. All the samples were measured in triplicate.

III. Result and discussion

1. Tracking α-glucosidase inhibitory constituents from *T. chebula* fruits with HPLC-based activity profiling

The inhibitory activity of fractions of *T. chebula* fruits against Baker's yeast α-glucosidase was shown in Table 24 .

Table 24. Inhibitory activity of the fractions of *T. chebula* fruits against Baker's yeast α-glucosidase

Fraction	IC_{50} (μg/ml)
n-Hexane fr.	>200
$CHCl_3$ fr.	132.3 ± 24.8
n-BuOH fr.	24.9 ± 4.7
B1 fr.	56.8 ± 9.4
B2 fr.	11.9 ± 1.6
B3 fr.	17.5 ± 1.6
B4 fr.	84.7 ± 9.8
Acarbose (positive control)	173.9 ± 25.6

As shown in Table 24, B2 and B3 fractions had the most significant inhibitory activity (IC_{50} 11.9 ± 1.6 and 17.5 ± 1.6 μg/ml, respectively). Thus, B2 and B3 was selected to perform HPLC-based activity profiling. To localize the active constituents, the B2 and B3 were microfractionated by semi-preparative HPLC into a 96-deep-well plates, which were lyophilized by freeze dryer, and residue in each microfraction was dissolved in buffer and assayed against Baker's yeast α-glucosidase. Figure 7 and 8 displayed the inhibitory effects of each microfraction. In this study, the inhibition hit limit was set at 20 %. In the B2 fraction, five

microfractions (23, 25, 26, 36, and 37) were found to be active. However, two microfractions (12 and 21) showed low inhibitory activity despite the large peak area. In the B3 fraction, six microfractions (12, 13, 20, 26, 39, and 41) were found to be active. To achieve preliminary information of each bioactive well, dereplication was performed by UHPLC-qTOF-MS and to verify the research method, further studies (dereplication and targeted isolation) were also performed to two non-bioactive microfractions.

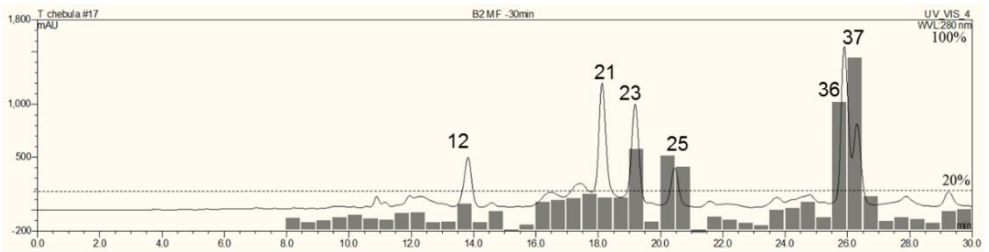


Figure 7. HPLC-based acitivity profiling of B2 fraction

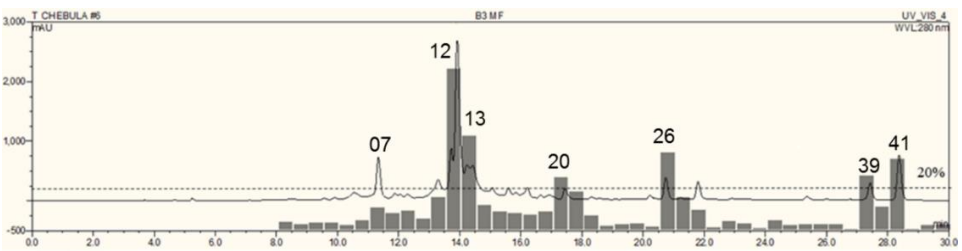


Figure 8. HPLC-based acitivity profiling of B3 fraction

Dereplication of these microfractions was performed based on their UV, HRMS, and MS/MS fragment. In addition, following spectral features of UV spectrum and MS/MS fragment were used for dereplication (Hanhineva et al., 2008; Pfundstein et al., 2010; Yang et al., 2012).

UV/Vis spectra:

1) Simple gallotannins have characteristic UV-Vis spectra with a maximum at

220 and 280 nm.

2) Ellagitannins have UV-Vis spectra with maximum absorbance below 275 nm.

3) The more HHDP unit presents, the higher is the relative intensity at 240-250 nm.

4) Ellagic acid unit containing compounds have characteristic UV/Vis spectra with a maximum at 254 and 365 nm.

MS/MS fragmentation:

1) Gallic acid esters loss of *m*/*z* 152 (galloyl) and 170 (gallate) from quasi-molecular ion $[M - H]^-$ and provide fragment ion at *m*/*z* 169 [Gallic acid ($C_7H_6O_5$) – H]$^-$.

2) Ellagitannins and ellagic acid glycosides loss of *m*/*z* 302 (ellagic acid) from quasi-molecular ion and provide fragment ion at *m*/*z* 301 [Ellagic acid ($C_{14}H_6O_8$) – H]$^-$ (Because free HHDP unit is simultaneously converted into ellagic acid).

3) Chebulic ellagitannins provide fragment ion at *m*/*z* 337 [Neochebuloyl ($C_{14}H_{12}O_{11}$) – H_2O – H]$^-$ or *m*/*z* 351 [Methyl neochebuloyl ($C_{15}H_{14}O_{11}$) – H_2O – H]$^-$.

4) 2,4-*O*-Chebuloyl moiety containing compounds loss of *m*/*z* 320 from quasi-molecular ion.

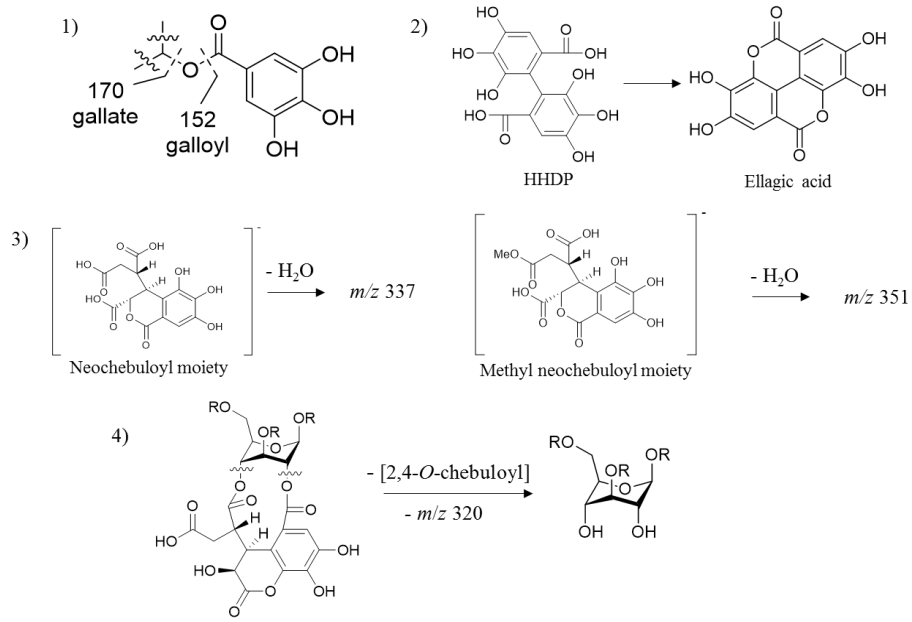


Figure 9. Features of MS fragmentation for dereplication of active microfractions.

Based on above features, thirteen compounds are tentatively identified as follows (Table 25 and Table 26). Especially, peak **12** and **13** were assumed as new compounds. The targeted isolation and inhibitory activity evaluation of the thirteen compounds were performed.

Table 25. Tentative identification of peaks from B2 fraction.

Microfraction No.	Peak No.	t_R	observed m/z (Δ ppm)	calculated m/z	Molecular formula [M-H]-	MS fragments	UV (λ_{max}, nm)	Tentative ID
B2-12	**1**	5.53	635.0905 (3.3)	635.0884	$C_{27}H_{23}O_{18}$	483, 465, 313, 295, 169	222, 278	Trigalloyl glucose
B2-21	**2**	6.70	953.0894 (-0.2)	953.0896	$C_{41}H_{29}O_{27}$	633, 463, 337, 301, 169	226, 275	Chebulagic acid
	3	7.00	985.1165 (0.7)	985.1158	$C_{42}H_{33}O_{28}$	633, 463, 351, 301, 169	226, 276	Methyl neochebulagate
B2-23	**4**	6.80	987.1324 (0.9)	985.1315	$C_{42}H_{35}O_{28}$	635, 483, 465, 351, 169	226, 280	Methyl neochebulinate
B2-25	**5**	7.10	787.0992 (-0.3)	787.0994	$C_{34}H_{27}O_{22}$	635, 617, 465, 447, 169	221, 278	Tetragalloyl glucose
	6	7.16	787.0990 (-0.5)	787.0994	$C_{34}H_{27}O_{22}$	635, 617, 465, 447, 169	221, 278	Tetragalloyl glucose
B2-36	**7**	7.86	955.1058 (0.5)	955.1053	$C_{41}H_{31}O_{27}$	785, 767, 635, 465, 337, 169	228, 279	Chebulinic acid
B2-37	**8**	8.27	939.1124 (2.1)	939.1104	$C_{41}H_{31}O_{26}$¡¡	787, 769, 617, 599, 447, 169	226, 280¡¡	¡¡1,2,3,4,6-*O*-Pentagalloyl glucose

Table 26. Tentative identification of peaks from B3 fraction.

Microfraction No.	Peak No.	t_R	observed m/z (Δ ppm)	calculated m/z	Molecular formula [M-H]-	MS fragments	UV (λ_{max}, nm)	Tentative ID
B3-12	**7**	7.86	955.1077 (2.5)	955.1053	$C_{41}H_{31}O_{27}$	785, 767, 617, 465, 337, 169	228, 279	Chebulinic acid
B3-13	**8**	8.28	939.1119 (1.6)	939.1104	$C_{41}H_{31}O_{26}$	787, 769, 617, 599, 447, 169	226, 280	1,2,3,4,6-*O*-Pentagalloyl glucose
B3-20	**9**	9.46	599.0679 (1.0)	599.0673	$C_{27}H_{19}O_{16}$	447, 301, 169	254, 361	4-*O*-(4"-*O*-Galloyl-*α*-L-rhamnosyl)ellagic acid
B2-25	**10**	11.1	751.0765 (-2.4)	751.0783	$C_{34}H_{23}O_{20}$	599, 449, 300, 169	254, 361	4-*O*-(2″,4″-Di-*O*-galloyl-*α*-l-rhamnosyl)ellagic acid and its isomer
	11	11.3	751.0795 (1.6)	751.0783	$C_{34}H_{23}O_{20}$	599, 449, 300, 169	255, 361	
B3-39	**12**	15.4	917.1437 (2.6)	917.1413	$C_{41}H_{33}O_{23}$	765, 747, 595, 579, 447, 169	223, 281	Unknown
B3-41	**13**	15.6	765.1323 (2.6)	765.1303	$C_{36}H_{29}O_{19}$	613, 595, 443, 425, 169	223, 279	Unknown

2. Elucidation of chemical structures of isolated compounds from *T. chebula* with HPLC-based activity profiling.

2.1. Compound **1**

Compound **1** was obtained as white amorphous powder. The negative HRESIMS spectrum of **1** gave an $[M - H]^-$ ion at *m*/*z* 635.0982 (calcd for $C_{27}H_{24}O_{18}$, 635.0884), indicating a molecular formula of $C_{27}H_{24}O_{18}$. The ^1H NMR showed signals for three galloyl groups [δ_H, 7.16, 7.13, 7.09 (each 2H, s), Gal-H-2″, 6″, Gal-H-2′, 6′, Gal-H-2‴, 6‴]. The presence of *β*-glucopyranose was deduced from large coupling constants of glucose H-1 (*d*, *J* = 8.2 Hz) signal (Figure 10). The locations of the three galloyl groups in the glucose moiety were determined to be the C-1, 3, and 6 positions by strongly downfielded ^1H NMR signals of H-1 (δ_H 5.83), H-3 (δ_H 5.27), and H-6 (δ_H 4.58 and 4.45). From obtained spectral data and above analysis, compound **1** was identified as 1,3,6-tri-*O*-galloyl-*β*-D-glucose (Zhang et al., 2004).

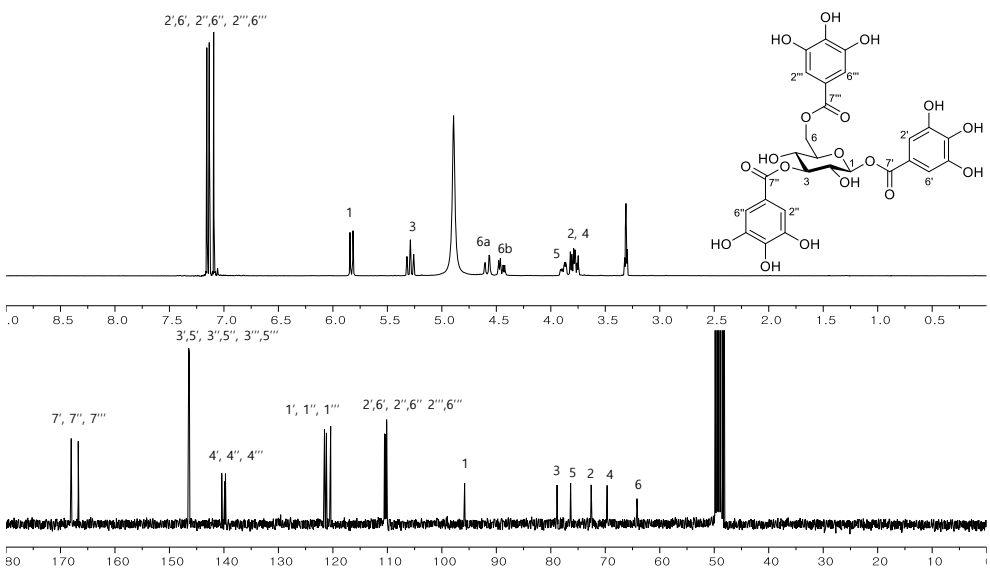


Figure 10. ^1H and ^{13}C NMR spectra of compound **1**

2.2. Compound **2**

Compound **2** was obtained as yellowish amorphous powder. Its molecular formula was identified as $C_{41}H_{30}O_{27}$ by negative HRESIMS spectrum (*m*/*z* 953.0865 $[M - H]^-$, calcd for $C_{41}H_{29}O_{27}$, 953.0896). The $^1C_4$ conformation of the β-glucopyranose was deduced from the small coupling constants of the H-1 – H-4 [δ_H 6.51 (1H, *s*, H-1), 5.83 (1H, *s*, H-3), 5.39 (1H, *s*, H-2), 5.23 (1H, *d*, *J* = 3.6 Hz, H-4)] signals and carbon signal of C-1 at δ_C 92.5 (Nawwar et al., 1994) (Figure 11). Its ^{1}H NMR spectrum showed that the presence of one galloyl group [δ_H 7.01 (2H, *s*, Gal-H-2′, 6′)] and one HHDP group [δ_H 6.84 (1H, *s*, HHDP-H-5′), 6.63 (1H, *s*, HHDP-H-5″)]. Furthermore, presence of a chebuloyl moiety was confirmed from one methylene at δ_H 2.14 (2H, *m*, Cheb-H-5′) and three methines at δ_H 5.05 (1H, *dd*, *J* = 1.2, 7.3 Hz, Cheb-H-3′), 4.80 (1H, *m*, Cheb-H-2′) and 3.90 (1H, *ddd*, *J* = 1.1, 3.6, 11.7 Hz, Cheb-H-4′) and one-proton singlet at δ_H 7.48 (1H, *s*, Cheb-H-2″). The galloyl and HHDP group were determined to be attached to C-1 and C-3/C-6 of core glucose, respectively, by HMBC cross peak between H-1 (δ_H 6.51) and galloyl carbonyl carbon (δ_C 166.4) and H-3 (δ_H 5.90)/H-6 (δ_H 4.78 and 4.41) and HHDP carbonyl carbon signals (δ_C 167.5 and 170.1). Furthermore, the correlation between H-2 (δ_H 5.39) and H-4 (δ_H 5.23) with chebuloyl carbonyl carbon (C-7′ and 7″, δ_C 166.2 and 174.4) showed that the chebuloyl unit was attached to C-2 and C-4 of the central glucose. Thus, compound **2** was identified as chebulagic acid (Pfundstein et al., 2010).

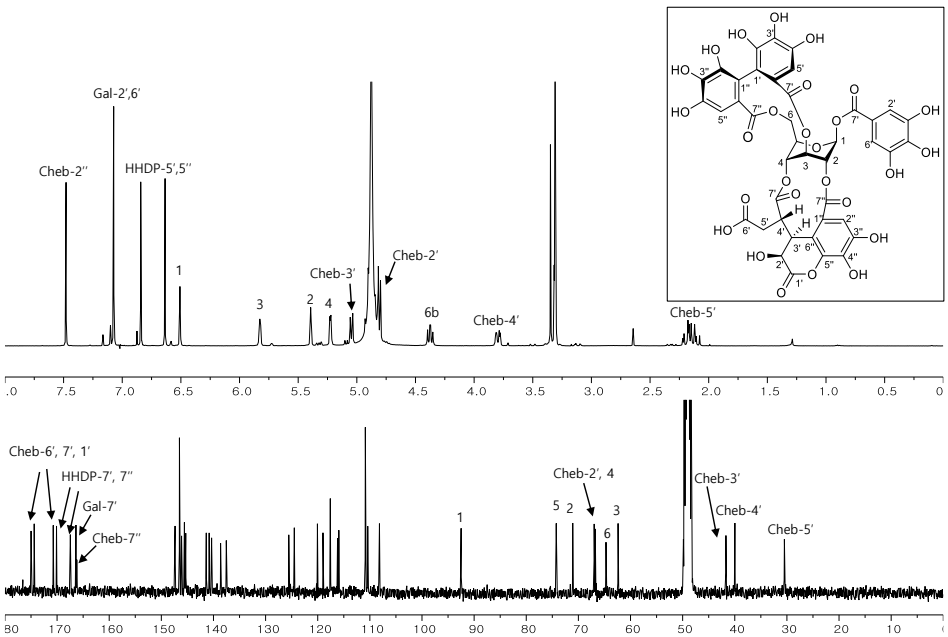


Figure 11. ^{1}H and ^{13}C NMR spectra of compound **2**

2.3. Compound **3**

Compound **3** was obtained as tan amorphous powder and had the molecular formula $C_{42}H_{34}O_{28}$ as determined by the HRESIMS at m/z 985.1151 $[M - H]^-$ (calcd. for $C_{42}H_{33}O_{28}$, 985.1158). The ^{1}H NMR spectrum of **3** showed signals at δ_H 7.10 (2H, *s*, Gal-H-2′, 6′), 6.83 (1H, *s*, HHDP-H-5′), 6.70 (1H, *s*, HHDP-5″) indicating the presence of a galloyl and an HHDP unit. The *β*-glucopyranose with $^{1}C_4$ conformation was deduced from seven signals with a discrete ^{1}H-^{1}H COSY system and the coupling constants (3.1–4.0 Hz) of the H-1–H-4 signals and carbon signal of C-1 at δ_C 94.7. The ^{1}H–^{1}H COSY spectrum also showed consecutive correlations from signal at δ_H 5.42 (1H, *d*, *J* = 0.7 Hz, Neocheb-H-2′) to δ_H 2.83 (1H, *dd*, *J* = 9.0, 17.3 Hz, Neocheb-H-5′a) and 2.51 (1H, *dd*, *J* = 5.1, 17.3 Hz, Neocheb-H-5′b) through δ_H 3.98 (1H, *d*, *J* = 9.1 Hz, Neocheb-H-3′) and δ_H 3.27 (1H, *td*, *J* = 9.1, 5.1 Hz, Neocheb-H-4′) in the neochebuloyl unit (Figure 12). Except for methoxy protons (δ_H 3.52), the spectroscopic data were very similar to

those of neochebulagic acid (**22**) (Saijo et al., 1989). HMBC correlations were observed between carboxylic carbon signals at δ_C 167.1 (HHDP-C-7′) and 168.8 (HHDP-C-7″) and the HHDP protons at δ_H 6.83 (HHDP-H-5′) and 6.70 (HHDP-5″) as well as between the carboxylic carbon signals at δ_C 165.8 (Gal-C-7′) and the galloyl H-2′, 6′ proton resonances at δ_H 7.10 ppm. The H-3 (δ_H 4.90) and H-6 (δ_H 4.59 and 4.29) signals were correlated to the HHDP carboxylic carbon signals at δ_C 167.1 and 168.8 ppm in the HMBC spectrum, respectively, and indicated that the HHDP unit was connected to C-3 and C-6 of the central glucose. Similarly, it was confirmed that the galloyl unit was connected to C-1 of the glucose. The correlation between H-4 (δ_H 5.67) with neochebuloyl carboxylic carbon at δ_C 172.6 (Neocheb-C-7′) showed that the neochebuloyl unit was connected to C-4 of the central glucose. The methoxy protons of **3** (δ_H 3.52) was correlated to C-6′ of the neochebuloyl unit (δ_C 172.8) in its HMBC spectrum (Figure 14). The absolute configuration of HHDP group was determined to be (*R*)-configuration by CD spectrum of corilagin (**42**), which was isolated from a partial hydrolysate of **3**, due to a strong negative Cotton effect at 237 nm (Okuda et al., 1982). To determine the absolute configuration of the neochebuloyl unit, compound **3** underwent acid hydrolysis and yielded chebulic acid. The absolute configuration of chebulic acid (**34**) was determined as (2′*S*, 3′*S*, 4′*S*) according to a positive Cotton effect at 235 nm in the CD spectrum (Yoshida et al., 1982). In addition, acid hydrolysate of **3** yielded D-glucose which was confirmed by HPLC analysis of its thiocarbamoyl thiazolidine derivative (Tanaka et al., 2007). Thus, compound **3** was characterized as 6′-*O*-methyl neochebulagate and it was first isolated from nature.

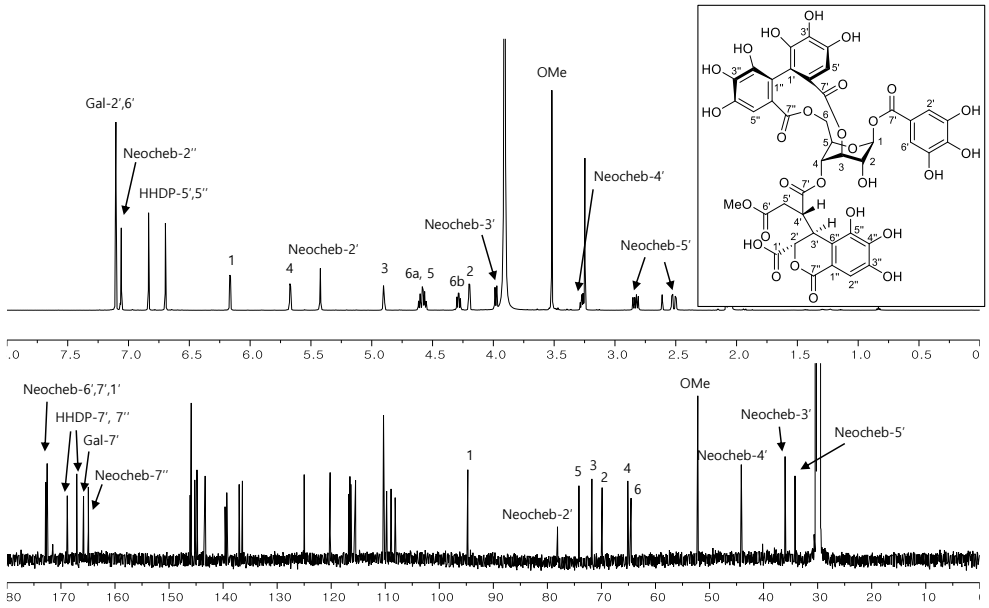


Figure 12. ^{1}H and ^{13}C NMR spectra of compound **3**

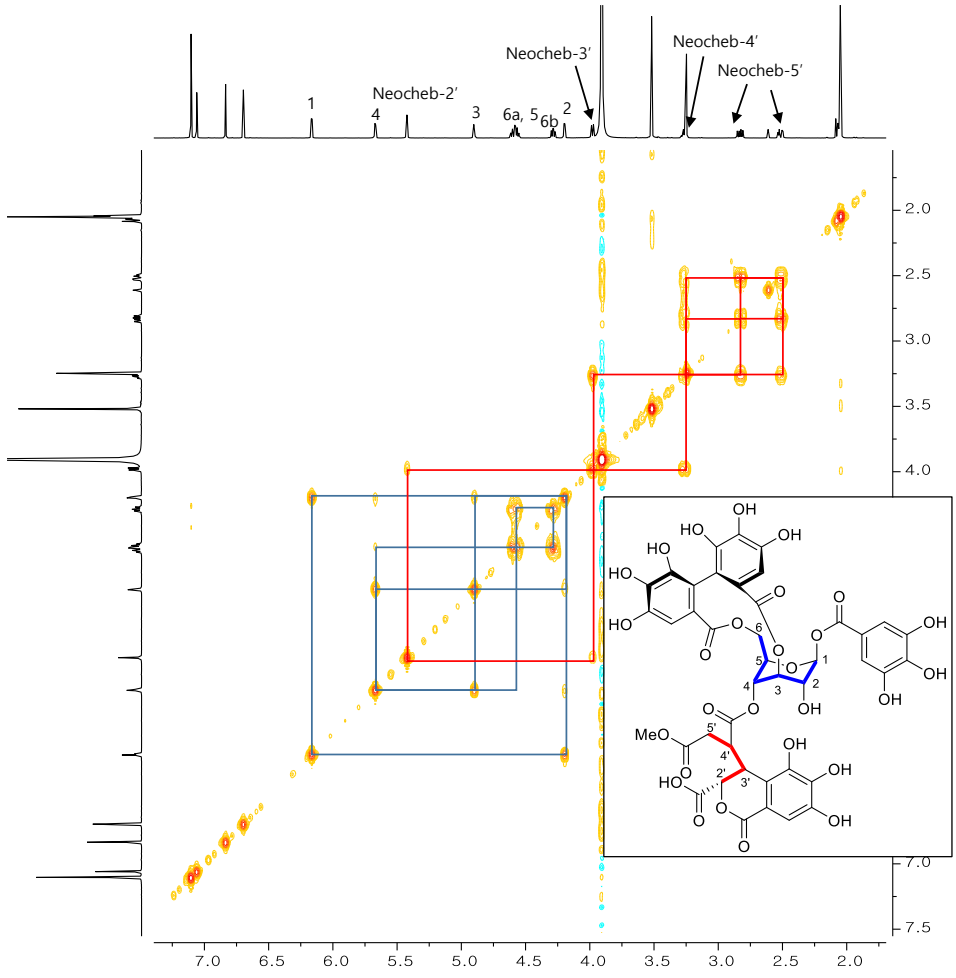


Figure 13. ^{1}H-^{1}H COSY spectrum of compound **3**

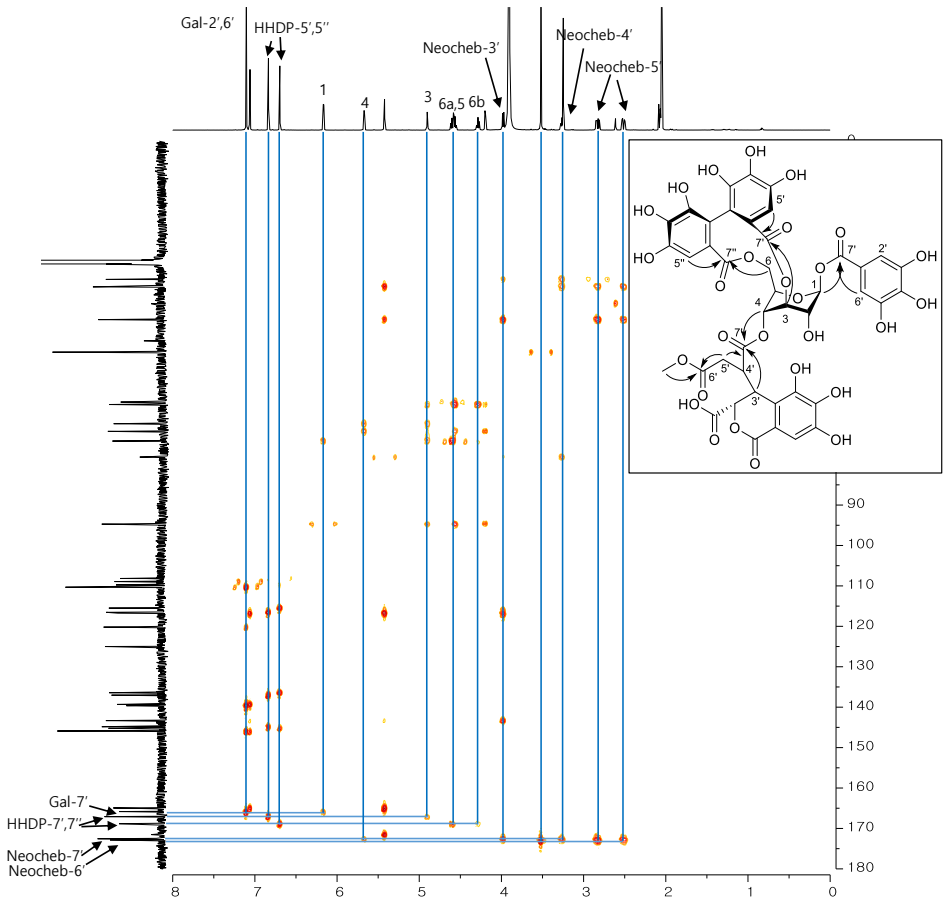


Figure 14. HMBC spectrum of compound **3**

2.4. Compound **4**

Compound **4**, a yellowish amorphous powder, was formulated as $C_{42}H_{36}O_{28}$ from the (-)-HRESIMS (*m/z* 987.1348 $[M - H]^-$, calcd for $C_{42}H_{35}O_{28}$, 987.1315). Its ^1H NMR spectra showed signals at δ_H 7.16 (2H, *s*, Gal-H-2′,6′), and 7.15 (4H, *s*, Gal-H-2″,6″, Gal-H-2‴,6‴), which suggested the presence of three galloyl groups. The existence of *β*-glucopyranose with $^4C_1$ conformation was confirmed from the large coupling constants (8.4–9.7 Hz) of the H-1 – H-4 signals [δ_H 6.02 (1H, *d*, *J* = 8.4 Hz, H-1), 5.66 (1H, *t*, *J* = 9.5 Hz, H-3), 5.53 (1H, *t*, *J* = 9.7 Hz, H-4), 3.99 (1H, *m*, H-2)]. Furthermore, three methines at δ_H 4.80 (1H, *d*, *J* = 1.2 Hz, Neocheb-H-2′),

3.73 (1H, *dd*, *J* = 1.6, 11.2 Hz, Neocheb-H-3′), 3.15 (1H, *ddd*, *J* = 4.3, 9.9, 11.2 Hz, Neocheb-H-4″), one methylene at 2.86 (1H, *dd*, *J* = 10.0, 17.7 Hz, Neocheb-H-5′a), 2.37 (1H, *dd*, *J* = 4.3, 17.7 Hz, Neocheb-H-5′b), and one-proton aromatic singlet at δ_H 7.05 (1H, *s*, Neocheb-H-2″) indicated the presence of neochebuloyl moiety (Figure 15). Except for the methoxy protons [δ_H 3.28 (3H, *s*, OMe)], the chemical shifts and coupling constants of these signals were analogous to neochebulinic acid (Qi et al., 2013). The HMBC confirmed that the three galloyl and one neochebuloyl units were attached to C-1 (δ_C 95.2), 3 (δ_C 75.9), 6 (δ_C 62.8) and C-4 (δ_C 70.0) of glucose moiety, respectively. The methoxy protons were correlated to Neocheb-C-1′ (δ_C 169.7) in the HMBC spectrum. With above observed spectroscopic data, compound **4** was determined to be 1′-*O*-methyl neochebulinate (Pfundstein et al., 2010).

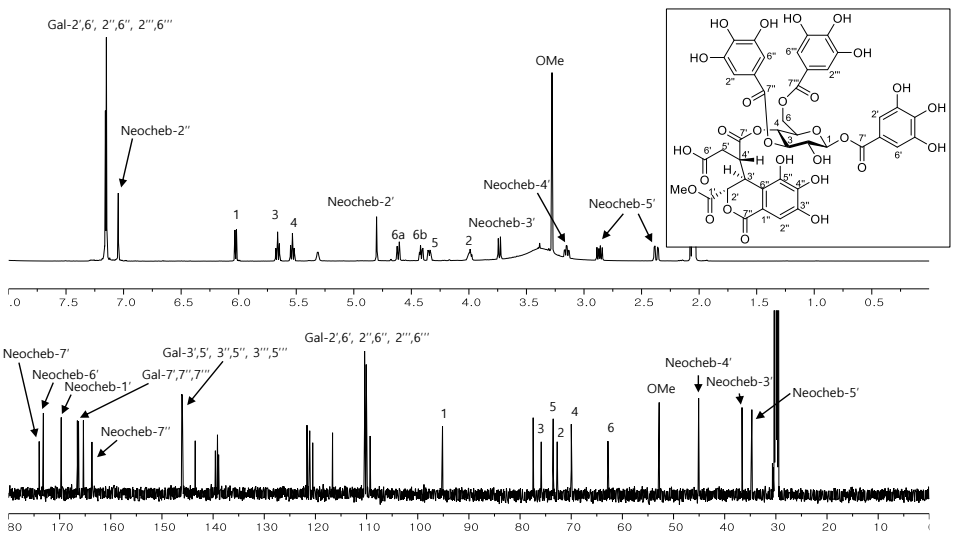


Figure 15. ^{1}H and ^{13}C NMR spectra of compound **4**

2.5. Compounds **5** and **6**

Compound **5** was obtained as white amorphous powder. The negative HRESIMS exhibited quasimolecular ion peak at *m*/*z* 787.1005 $[M - H]^-$ (calcd for $C_{34}H_{27}O_{22}$,

787.0994) indicating $C_{34}H_{28}O_{22}$ as its molecular formula. The ^{1}H NMR spectroscopic data displayed for four two-proton aromatic singlets [δ_H, 7.15, 7.07, 7.05, 6.97 (each 2H, s), Gal-H-2″, 6″, Gal-H-2′, 6′, Gal-H-2‴, 6‴, Gal-H-2⁗, 6⁗] assignable to four galloyl units. Additionally, the anomeric proton signal at δ_H, 6.16 (1H, *d*, *J* = 8.4 Hz, H-1) suggested the presence of *β*-glucopyranose (Figure 16). The locations of four galloyl moieties in the glucose unit were determined to be the C-1, 2, 3, and 6 positions by strongly downfielded ^{1}H NMR signals of H-1 (δ_H 6.16), H-2 (δ_H 5.45), H-3 (δ_H 5.66), and H-6 (δ_H 4.56). With above mentioned spectroscopic data, compound **5** was determined as 1,2,3,6-tetra-O-galloyl-*β*-D-glucose (Tanaka et al., 1983).

Compound **6** was obtained as white amorphous powder and showed the ion peak at *m*/*z* 787.1002 [M – H]$^-$ (calcd for $C_{34}H_{27}O_{22}$, 787.0994) in the negative HRESIMS indicating its molecular formula ($C_{34}H_{28}O_{22}$). Like compound **5**, compound **6** had four two-proton aromatic signals [δ_H, 7.20, 7.15, 7.06, 6.04 (each 2H, s), Gal-H-2″, 6″, Gal-H-2′, 6′, Gal-H-2‴, 6‴, Gal-H-2⁗, 6⁗] and one anomeric proton signal [δ_H 6.04 (1H, *d*, *J* = 8.4 Hz, H-1)]. These characteristic ^{1}H NMR data suggested that compounds **6** was also tetragalloylglucose (Figure 17). In comparison with compound **5**, the ^{1}H NMR signals of H-2 (δ_H 4.04) moved to upfield and H-4 (δ_H 5.49) moved to downfield. On the basis of above information, compound **6** was determined as 1,3,4,6-tetra-O-galloyl-*β*-D-glucose (Tanaka et al., 2003).

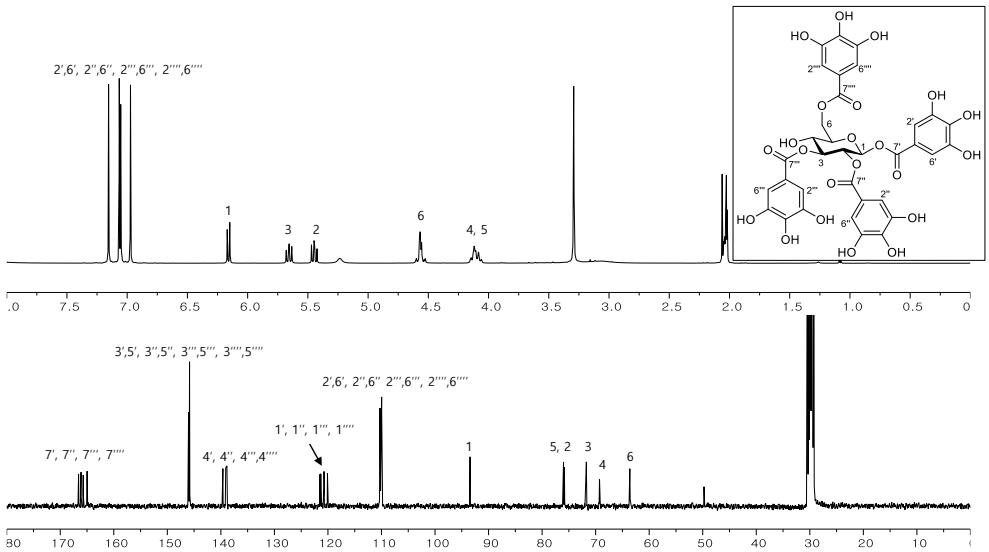


Figure 16. ^{1}H and ^{13}C NMR spectra of compound **5**

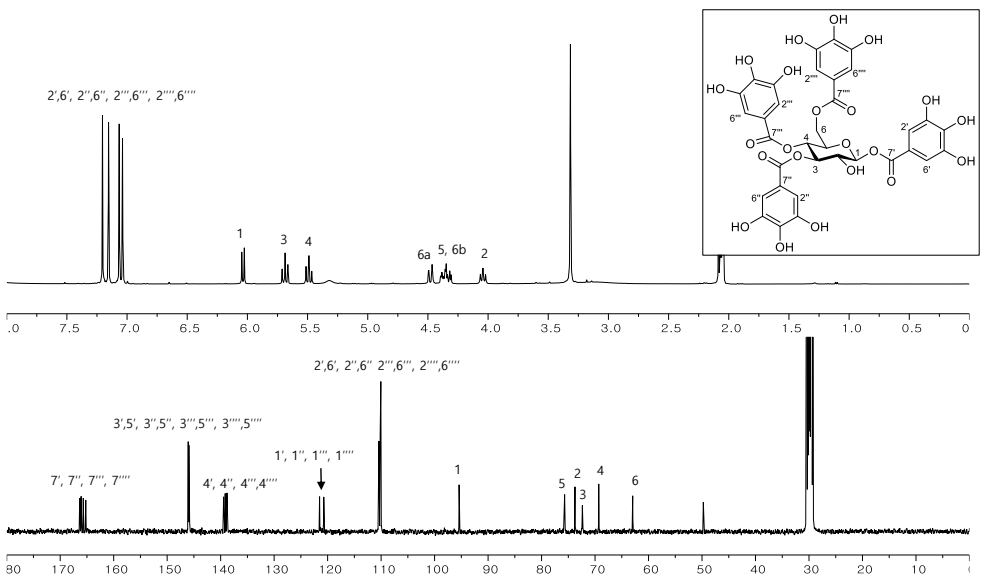


Figure 17. ^{1}H and ^{13}C NMR spectra of compound **6**

2.6. Compound **7**

Compound **7** was obtained as white amorphous powder. The negative HRESIMS spectrum of **7** showed the $[M - H]^-$ peak at *m/z* 955.1057 (calcd for $C_{41}H_{31}O_{27}$,

955.1053) and established the molecular formula $C_{41}H_{32}O_{27}$. The ^{1}H NMR spectrum showed characteristic signals for three galloyl moieties at δ_H 7.17 (2H, *s*, Gal-H-2″, 6″), 7.10 (2H, *s*, Gal-H-2′, 6′), 6.98 (2H, *s*, Gal-2‴, 6‴). The presence of a glucose moiety was apparent from the ^{1}H NMR signals at δ_H 6.49 (1H, *d*, *J* = 2.8 Hz, H-1) 6.24 (1H, br *s*, H-3), 5.43 (1H, *d*, *J* = 1.2 Hz, H-2), 5.04 (1H, *d*, *J* = 3.4 Hz, H-4), 4.81 (1H, *m*, H-6a), 4.72 (1H, *m*, H-5) 4.62 (1H, *dd*, *J* = 6.3 10.6 Hz, H-6b). Furthermore, the presence of one methylene at δ_H 2.24 (2H, *m*, Cheb-H-5′) and three methine protons at δ_H 5.09 (1H, *dd*, *J* = 1.4, 7.2 Hz, Cheb-H-3″), 4.82 (1H, *m*, Cheb-H-2′), 3.87 (1H, *ddd*, *J* = 0.9, 4.7, 10.5 Hz, Cheb-H-4′), and one aromatic proton at δ_H 7.51 (1H, s, Cheb-H-2″) on the ^{1}H NMR spectrum and carbon atoms of four carboxylic carbon signals at δ_C 175.1 (Cheb-C-6′), 174.6 (Cheb-C-7′), 170.1 (Cheb-C-1′) and 166.1 (Cheb-C-7″) on the ^{13}C NMR spectrum suggested the presence of chebuloyl group (Figure 18). Thus, compound **7** was confirmed as chebulinic acid by comparison with literature (Pfundstein et al., 2010).

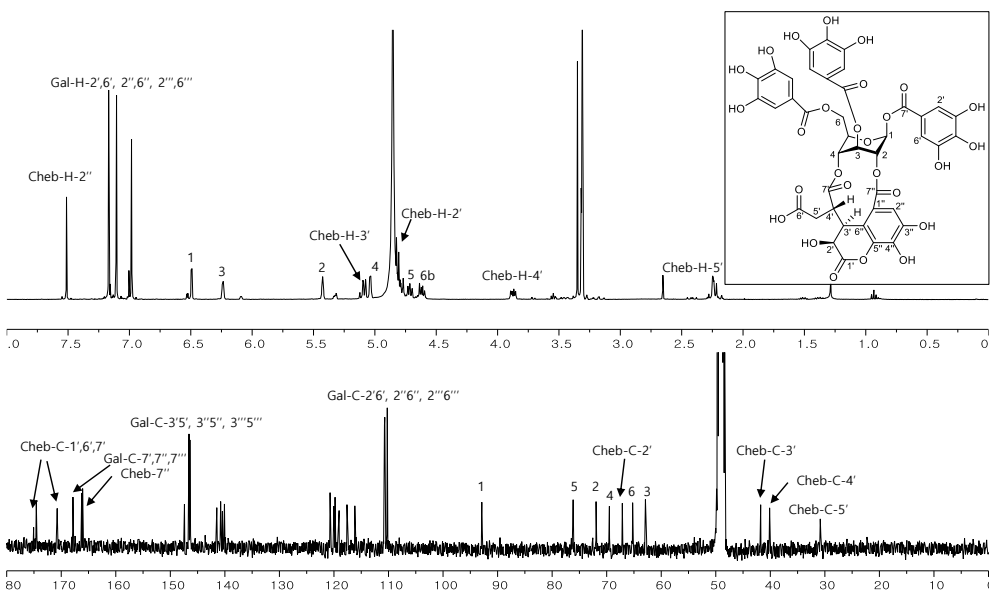


Figure 18. ^{1}H and ^{13}C NMR spectra of compound **7**

2.7. Compound **8**

Compound **8** was obtained as white amorphous powder. Its negative HRESIMS peak at *m*/*z* 939.1119 $[M - H]^-$ (calcd for $C_{41}H_{31}O_{26}$, 939.1104) established a molecular formula of $C_{41}H_{32}O_{26}$. The ^1H NMR spectrum exhibited signals for five galloyl units [δ_H, 6.97, 6.91, 6.84, 6.81, 6.76 (each 2H, s), Gal-H-2″, 6″, Gal-H-2′, 6′, Gal-H-2‴, 6‴, Gal-H-2⁗, 6⁗, Gal-H-2⁗′, 6⁗′] and one *β*-glucopyranose moiety [δ_H 6.37 (1H, *d*, *J* = 8.2 Hz, H-1)] (Figure 19). These ^1H NMR data suggested that compounds **8** was pentagalloylglucose. The locations of five galloyl units in the glucose moiety were determined to be the C-1, 2, 3, 4 and 6 positions by downfielded ^1H NMR signals of H-1 (δ_H 6.37), H-2 (δ_H 5.43), H-3 (δ_H 5.95), H-4 (δ_H 5.43), and H-6 (δ_H 4.30). With above mentioned spectroscopic data, compound **8** was determined as 1,2,3,4,6-penta-*O*-galloyl-*β*-D-glucose (Abdel-Mageed et al., 2014).

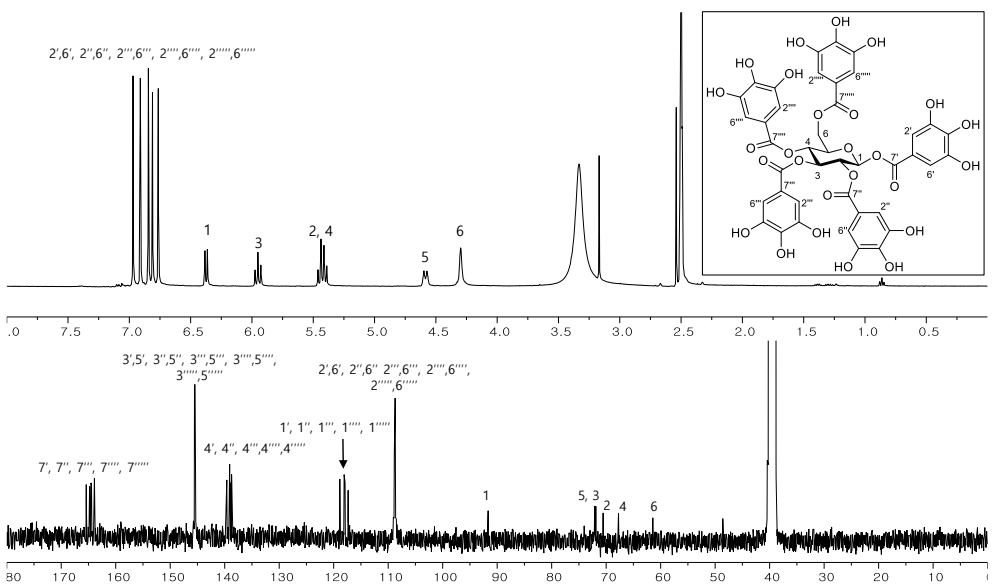


Figure 19. ^1H and ^{13}C NMR spectra of compound **8**

2.8. Compounds **9**-**11**

Compound **9** was obtained as brown amorphous powder, and $[\alpha]_D^{20}$= -149.0 (c 0.10, MeOH). The (-)-HRESIMS exhibited quasi-molecular ion peak at *m*/*z*: 599.0684 $[M - H]^-$ (calcd for $C_{27}H_{19}O_{16}$, 599.0673) indicating $C_{27}H_{20}O_{16}$ as its molecular formula. The presence of ellagic acid moiety was confirmed by its UV spectrum at 258 and 356 nm and two aromatic singlets at δ_H 7.89 and 7.51 (each 1H, H-5 and H-5′) (El-Toumy and Rauwald, 2002). The ^{1}H NMR spectrum also displayed two-proton aromatic singlet at δ_H 7.10 (2H, *s*, Gal-H-2′,6′) suggesting the presence of one galloyl unit. Furthermore, the presence of *α*-rhamnopyranose with $^1C_4$ conformation was deduced from five oxygenated methines [δ_H 5.65 (1H, *s*, H-1″), 5.25 (1H, *t*, *J* = 9.5 Hz, H-4″), 4.30 (1H, *m*, H-3″), 4.29 (1H, br *s*, H-2″), 4.04 (1H, *dq*, *J* = 12.4, 6.1 Hz, H-5″)] along with a doublet methyl proton at δ_H 1.19 (3H, *d*, *J* = 6.2 Hz, H-6″) and the very small vicinal coupling constants of the H-1″ and H-2″ signals (Figure 20). Rhamnosylation at C-4 of ellagic acid was confirmed by the HMBC correlation between the signals of H-1″ (δ_H 5.65) and C-4 (δ_C 147.7). The location of a galloyl unit in the rhamnose moiety was assigned to be the C-4″ position by downfielded ^{1}H NMR signals of H-4″ (δ_H 5.25) and HMBC correlation between the signals of H-4″ (δ_H 5.25) and Gal-C-7‴ (δ_C 168.1). Therefore, compound **9** was determined as 4-*O*-(4″-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (Pfundstein et al., 2010).

Compound **10** was obtained as brown amorphous powder, and $[\alpha]_D^{20}$= -60.5 (c 0.10, MeOH). Its molecular formula was confirmed as $C_{34}H_{24}O_{20}$ by (-)-HRESIMS (*m*/*z*: 751.0787 $[M - H]^-$, calcd for $C_{34}H_{23}O_{20}$, 751.0783) In the UV spectrum of **10**, absorption at 256 and 358 nm indicated the presence of ellagic acid moiety. Except for one additional galloyl moiety its ^{1}H and ^{13}C NMR spectra were analogous to those of compound **9**. The attachment of the one additional galloyl moiety in the rhamnose moiety was assigned to be the C-3″ position by downfielded ^{1}H NMR

signals of H-3″ (δ_H 5.66) and HMBC correlation between the signals of H-3″ (δ_H 5.66) and Gal-C-7‴ (δ_C 168.2). Therefore, compound **10** was confirmed as 4-*O*-(3″,4″-di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (Pfundstein et al., 2010).

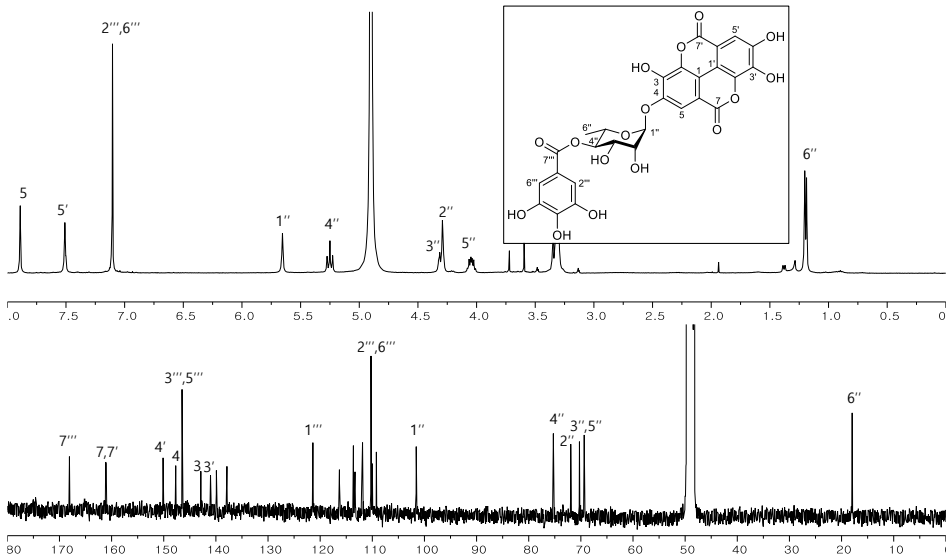


Figure 20. ^{1}H and ^{13}C NMR spectra of compound **9**

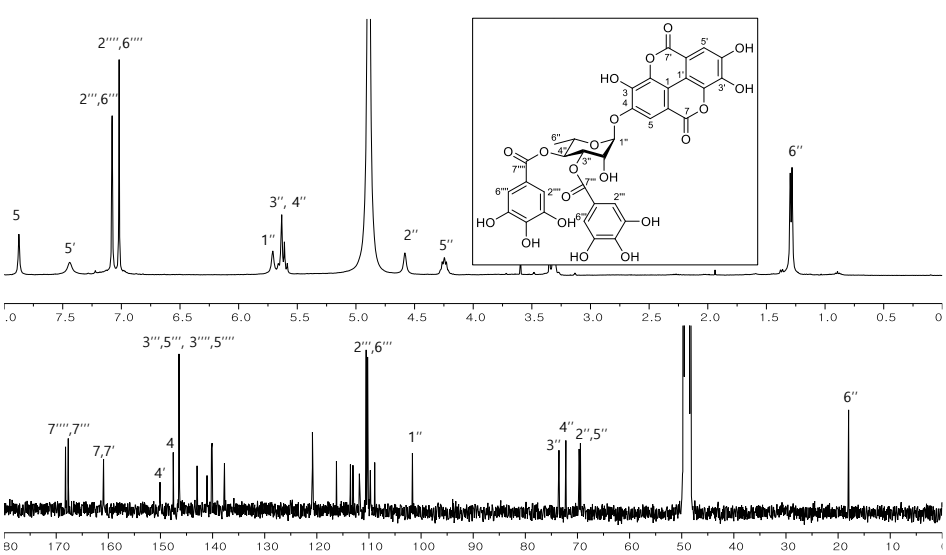


Figure 21. ^{1}H and ^{13}C NMR spectra of compound **10**

Compound **11** was obtained as brown amorphous powder, and $[\alpha]_D^{20}$= -37.4 (c 0.10, MeOH). It had the molecular formula $C_{34}H_{24}O_{20}$ as determined by the (-)-HRESIMS at *m*/*z*: 751.0786 $[M - H]^-$ (calcd for $C_{34}H_{23}O_{20}$, 751.0783). In the UV spectrum, absorption at 258 and 356 nm suggested that **11** also had ellagic acid moiety. Based on molecular formula, ^{1}H, and ^{13}C NMR spectra, compound **11** was regioisomer of **10**. Rhamnosylation at C-4 of ellagic acid was determined by the HMBC correlation between the signals of H-1″ (δ_H 5.76) and C-4 (δ_C 147.6). The H-4″ (δ_H 5.37) signal was correlated to galloyl carbonyl carbon signal at δ_C 168.0 (C-7‴) in the HMBC spectrum. However, despite of the presence of downfielded H-2″ signal at δ_H 5.65 (1H, br *s*), HMBC correlations in CD_3OD between H-2″ (δ_H 5.65) and remaining galloyl carbonyl carbon signal at δ_C 167.7 (C-7‴) was not observed (Figure 24). Therefore, we changed the solvent to DMSO-d_6 and observed the HMBC correlation between H-2″ and C-7‴ (Figure 25). The L-configuration of the rhamnose was identified by HPLC analysis of the thiocarbamoyl thiazolidine derivative of the sugar from acid hydrolysate of **11** compared with that of the authentic sample of L-rhamnose. Thus, compound **11** was characterized as 4-*O*-(2″,4″-di-*O*-galloyl-α-L-rhamnosyl)ellagic acid and was first isolated from nature.

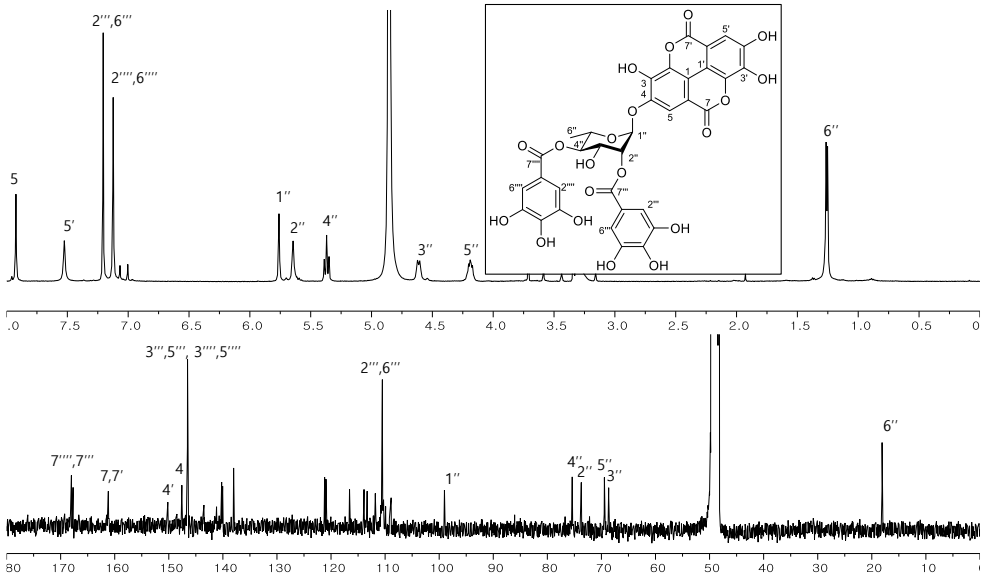


Figure 22. ^{1}H and ^{13}C NMR spectra of compound **11**

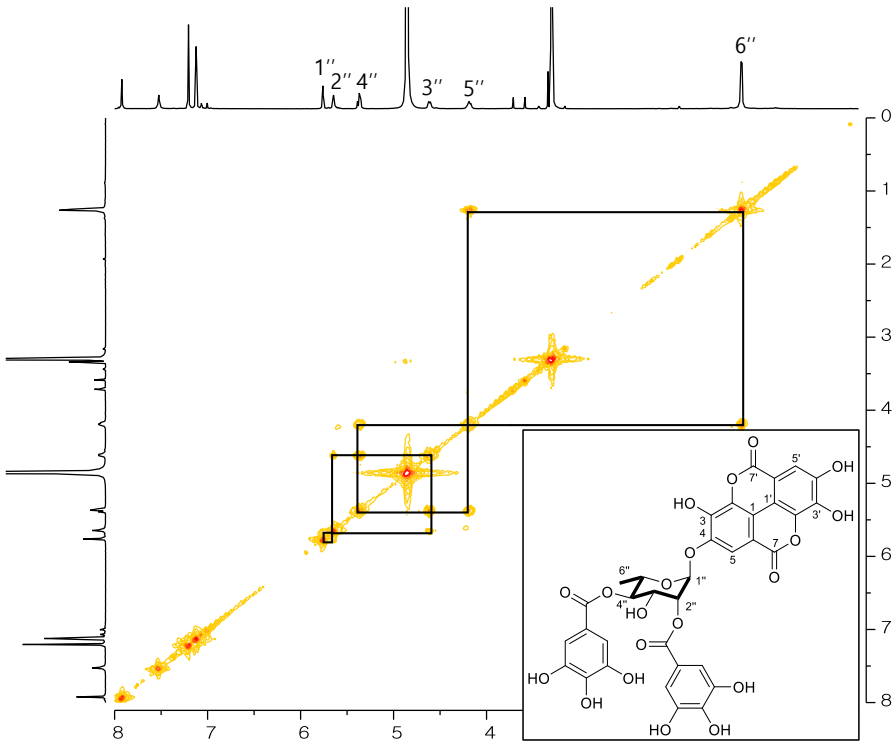


Figure 23. ^{1}H-^{1}H COSY spectrum of compound **11**

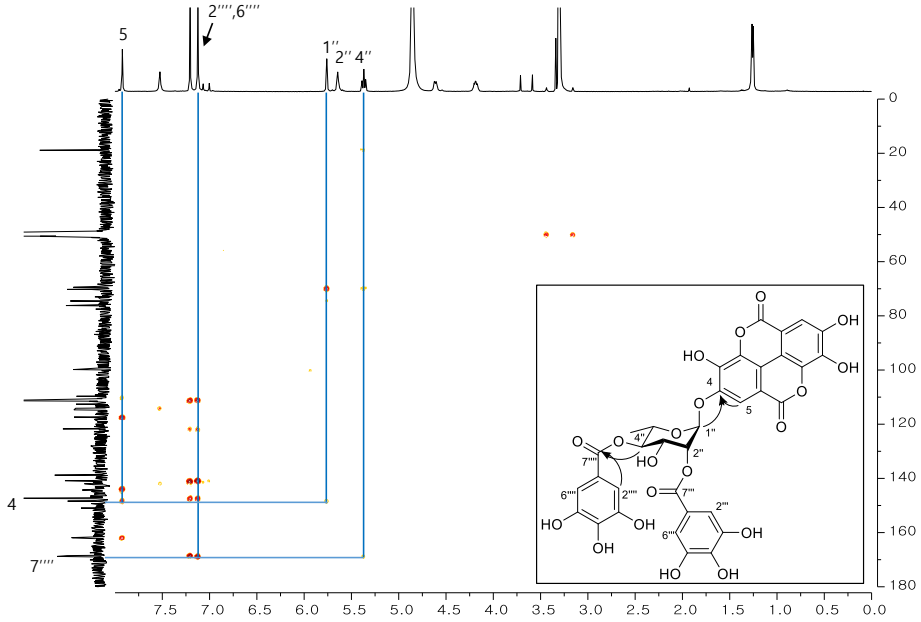


Figure 24. HMBC spectrum of compound **11** in CD_3OD

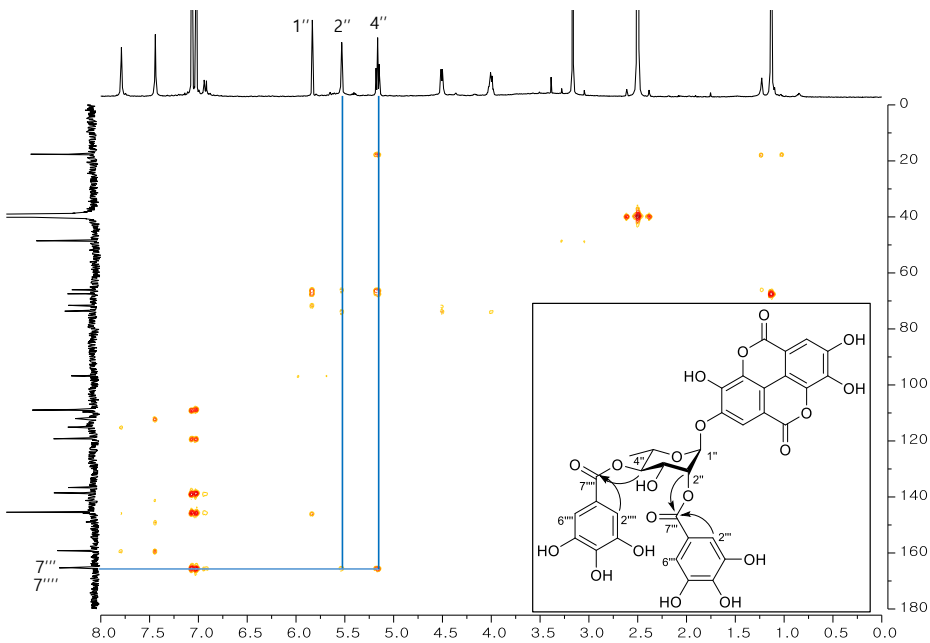


Figure 25. HMBC spectrum of compound **11** in DMSO-d_6

2.9. Compounds **12** and **13**

Compound **12** was obtained as brown amorphous powder and had the molecular formula $C_{43}H_{34}O_{23}$ as determined by HRESIMS at *m*/*z* 917.1425 $[M - H]^-$ (calcd for $C_{43}H_{33}O_{23}$, 917.1413). The ^{1}H-NMR spectrum showed four singlet signals at δ_H 7.13 (2H, s, Gal-H-2‴,6‴), 7.06 (Gal-H-2′,6′), 6.99, and 6.99 (each *s*, 2H, Gal-H-2″,6″, Gal-H-2⁗,6⁗), that suggested the presence of four galloyl units. In addition, *trans* olefinic signals at δ_H 7.63 (1H, *d*, *J* = 16.0 Hz, Cin-H-7′) and 6.45 (1H, *d*, *J* = 16.0 Hz, Cin-H-8′) and five aromatic proton multiplets (δ_H 7.71–7.43, Cin-H-1′ – Cin-H-5′) suggested that **12** had a cinnamoyl unit in its molecules (Kashiwada et al., 1988) (Figure 26). The *β*-glucopyranose unit was deduced from seven signals with a discrete ^{1}H–^{1}H COSY system and large coupling constants (8.2–9.4 Hz) of the glucose H-1 – H-4 signals (Era et al., 2015; Orabi et al., 2010). The H-1 (δ_H 6.25), H-3 (δ_H 5.94), and H-6 (δ_H 4.57 and 4.37) signals were correlated to the galloyl carbonyl carbon signals at 165.3 (Gal-C-7′), 166.4 (Gal-C-7‴) and 166.7 (Gal-C-7⁗) in the HMBC spectrum, respectively. However, the remaining galloyl and cinnamoyl carbonyl carbon signals were perfectly overlapped (δ_C 166.3), the connection of Glc H-2 (δ_H 5.60) and H-4 (δ_H 5.61) to the galloyl and cinnamoyl unit could not be distinguished in the HMBC spectrum (Figure 28). For this reason, we measured 1D and 2D NMR spectra of **2** in pyridine-d_5. As a result, four galloyl and a cinnamoyl carbonyl carbon signals were clearly separated. In the HMBC spectrum of **2** measured in pyridine-d_5, the H-2 (δ_H 6.22) signal was correlated to the galloyl carbonyl carbon signal at δ_C 166.8 (Gal-C-7‴) and the H-4 (δ_H 6.13) was correlated to the cinnamoyl carbonyl carbon signal at δ_C 166.2 (Cin-C-9′) which was correlated by olefinic proton signals at δ_H 7.76 (1H, *d*, *J* = 16.0 Hz, Cin-H-7′) and 6.56 (1H, *d*, *J* = 16.0 Hz, Cin-H-8′) of cinnamoyl moiety (Figure 29). Absolute configuration of sugar was identified to be D-glucose by RP-HPLC analysis of the thiocarbamoyl thiazolidine derivative of the sugar present in the

acid hydrolysate of **1** and that of authentic D-glucose (Tanaka et al., 2007). Thus, compound **12** was characterized as 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose and it was isolated for the first time from nature.

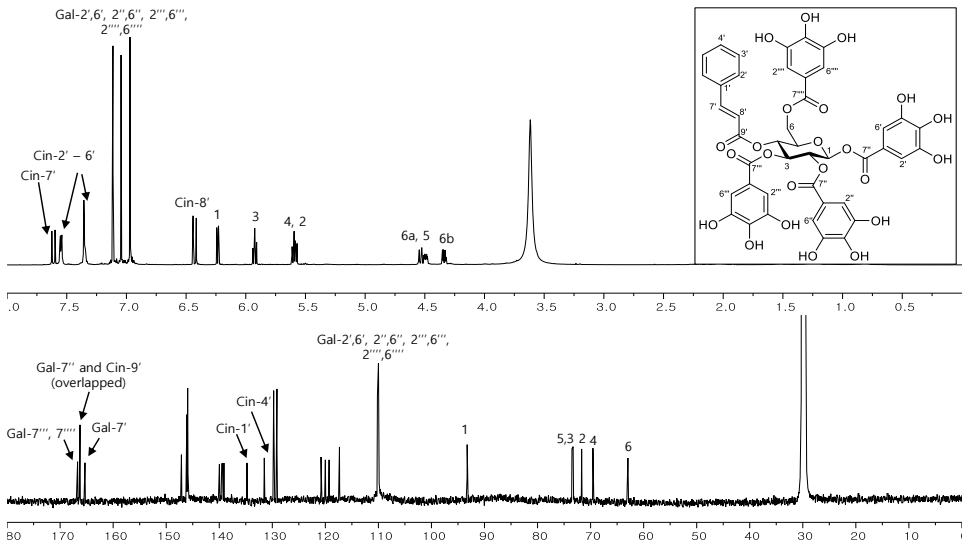


Figure 26. ^{1}H and ^{13}C NMR spectra of compound **12**

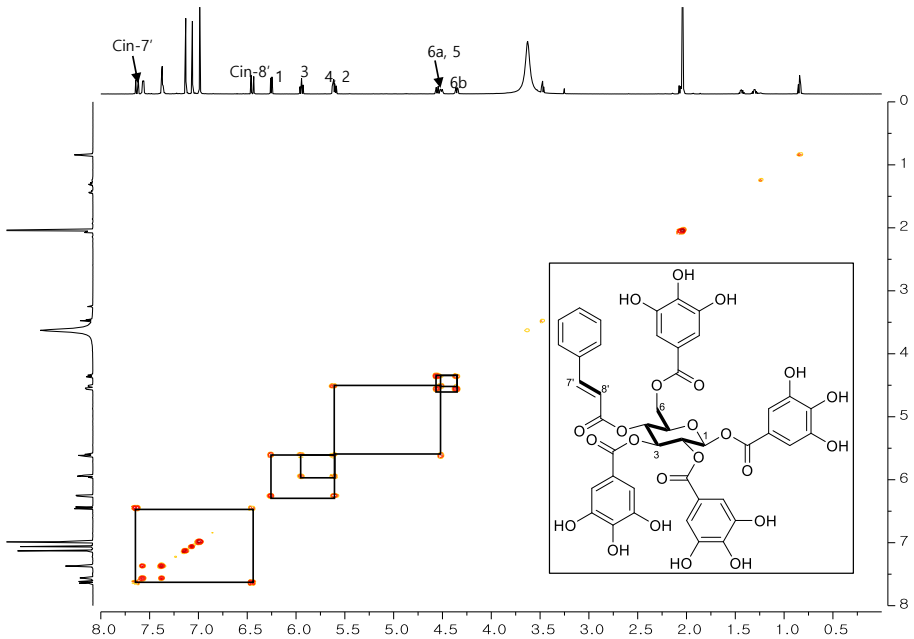


Figure 27. ^{1}H-^{1}H COSY spectrum of compound **12**

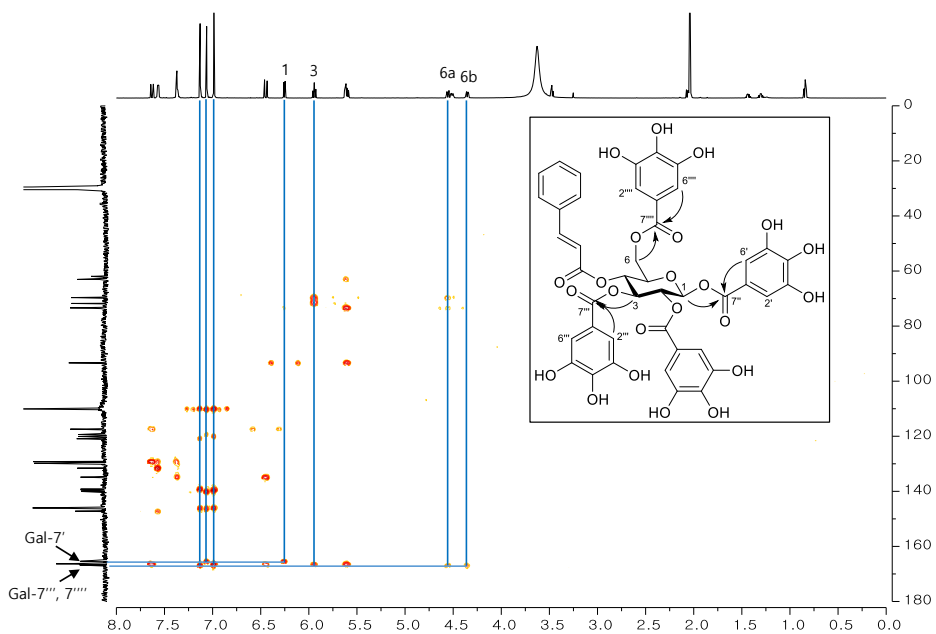


Figure 28. HMBC spectrum of compound **12** (solvent: acetone-d_6 : D_2O = 9 : 1)

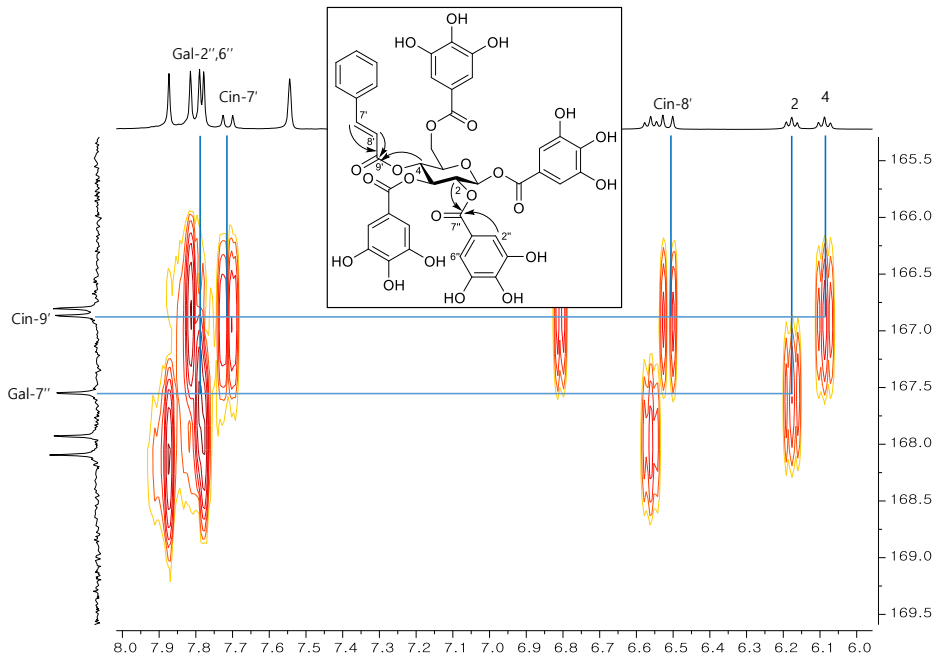


Figure 29. Expanded HMBC spectrum of compound **12** (solvent: pyridine-d_5)

Compound **13** was obtained as brown amorphous powder and had the molecular formula $C_{36}H_{30}O_{19}$ as determined by HRESIMS at *m*/*z* 763.1323 $[M - H]^-$ (calcd for $C_{36}H_{29}O_{19}$, 765.1303). Based on the ^{1}H NMR spectrum and difference in molecular formula between **12** and **13**, corresponding to $C_7H_4O_4$, **13** had one less galloyl unit than that of **12**. The H-1 (δ_H 6.18), 2 (δ_H 5.47), and 3 (δ_H 5.66) signals were correlated to the galloyl carbonyl carbon signals at δ_C 165.0 (Gal-C-7′), 165.8 (Gal-C-7″) and 166.2 (Gal-C-7‴) in the HMBC spectrum, respectively. The H-6 signals (δ_H 4.60 and 4.50) were correlated to the cinnamoyl carbonyl carbon signal at δ_C 167.1 in the HMBC spectrum (Figure 32). The presence of the D-glucose was confirmed by HPLC analysis of its thiocarbamoyl thiazolidine derivative. Thus, compound **13** was characterized as 1,2,3-tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose and newly isolated from nature.

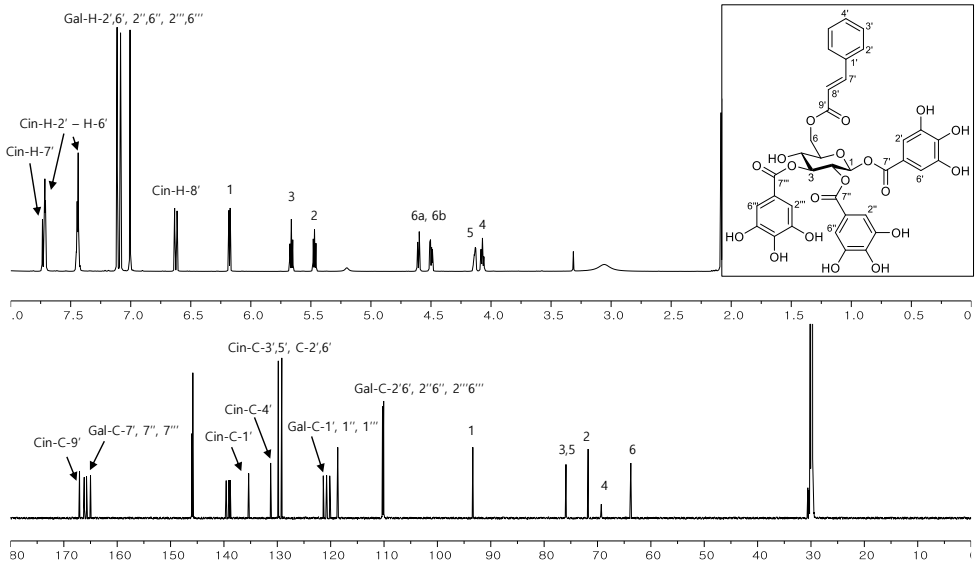


Figure 30. ^{1}H and ^{13}C NMR spectra of compound **13**

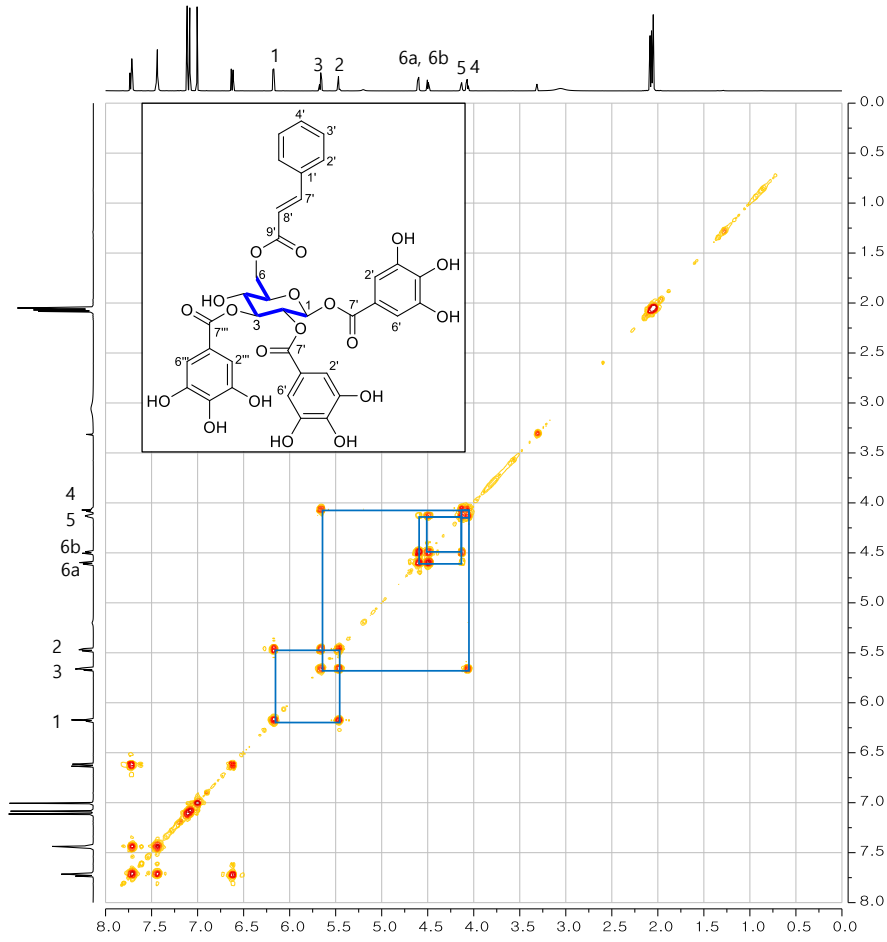


Figure 31. ^{1}H-1H COSY spectrum of compound **13**

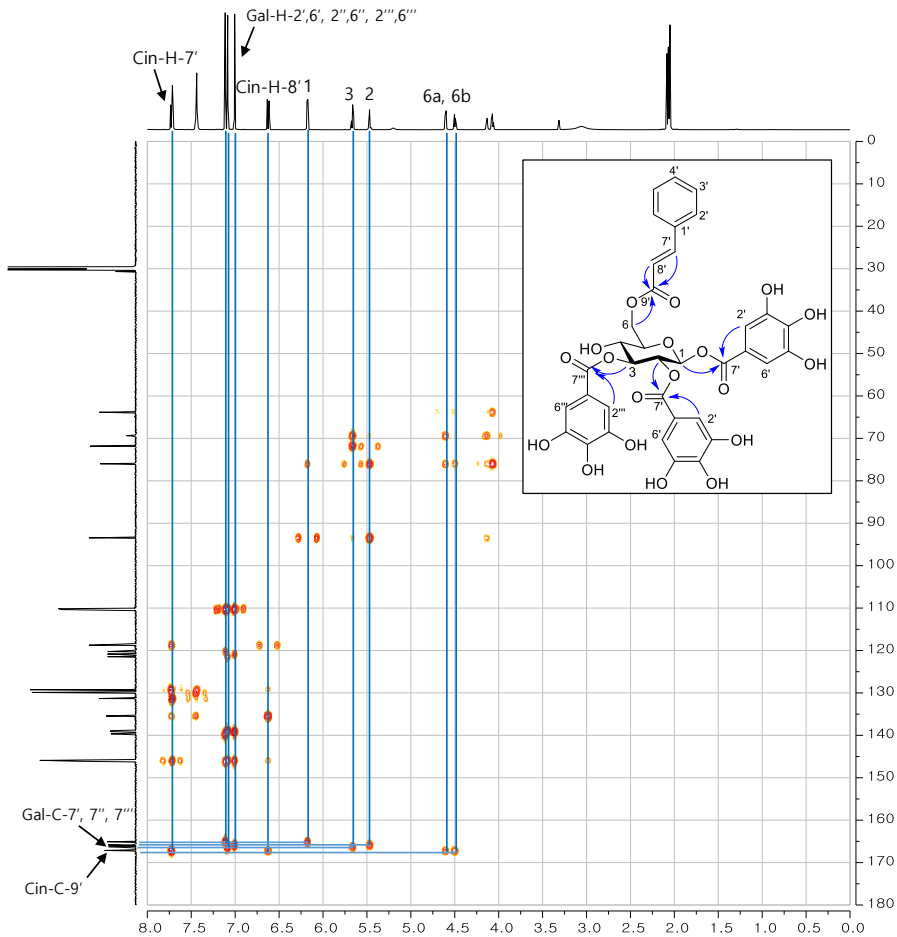


Figure 32. HMBC spectrum of compound **13**

These thirteen compounds were tested to evaluate their Baker's yeast α-glucosidase inhibitory activity (Table 27). Among these, all of the nine compounds (**4**–**13**) obtained from the active microfractions had inhibitory activity (IC_{50} 2.9 – 59.5 μM). On the other hand, all of the three compounds (**1**-**3**) isolated from non-bioactive microfractions had no inhibitory activity (IC_{50} > 200 μM). Based on these results, it is confirmed that HPLC-based activity profiling is an effective method for tracking targeted bioactive compounds.

Table 27. Baker's yeast α-glucosidase inhibitory activity of isolated compounds with HPLC-based activity profiling.

Compound	Name	IC_{50} (μM)
	Compounds from no bioactive microfraction	
1	1,3,6-Tri-*O*-galloyl-*β*-D-glucose	> 200
2	Chebulagic acid	> 200
3	6′-*O*-Methyl neochebulagate	> 200
	Compounds from bioactive microfraction	
4	1′-*O*-Methyl neochebulinate	59.5 ± 9.6
5	1,2,3,6-Tetra-*O*-galloyl-*β*-D-glucose	15.7 ± 0.8
6	1,3,4,6-Tetra-*O*-galloyl-*β*-D-glucose	15.5 ± 0.6
7	Chebulinic acid	24.0 ± 4.0
8	1,2,3,4,6-Penta-*O*-galloyl-*β*-D-glucose	8.3 ± 0.5
9	4-*O*-(4″-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid	37.9 ± 6.7
10	4-*O*-(3″,4″-Di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid	12.3 ± 2.0
11	4-*O*-(2″,4″-Di-*O*-galloyl-α-L-rhamnosyl)ellagic acid	6.4 ± 1.0
12	1,2,3,6-Tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose	2.9 ± 0.8
13	1,2,3-Tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose	14.7 ± 2.7
Positive control	Acarbose	174.0 ± 29.8

Additionally, to investigate constituents of *T. chebula* fruits, further isolation was performed.

3. Elucidation of chemical structures of further isolated compounds from *T. chebula*.

3.1. Compounds **14** and **15**

Compound **14** was obtained as white amorphous powder and yielded a molecular formula of $C_{20}H_{20}O_{14}$ on the negative HRESIMS peak at *m*/*z* 483.0777 $[M - H]^-$ (calcd for $C_{20}H_{19}O_{14}$, 483.0775) and its ^{13}C NMR. Analyzing its 1H NMR data (Figure 33) revealed that its structure had two galloyl moieties [δ_H, 7.14, 7.12, (each 2H, s), Gal-H-2′, 6′, Gal-H-2″, 6″] and one *β*-glucopyranose [δ_H, 5.79 (1H, *d*, *J* = 8.2), H-1]. The locations of two galloyl units in the glucose moiety were assigned to be the C-1 and 3 positions by downfielded ^1H NMR signals of H-1 (δ_H 5.79) and H-3 (δ_H 5.24). On the basis of these data, compound **14** was determined as 1,3-di-*O*-galloyl-*β*-D-glucose (Li et al., 2015).

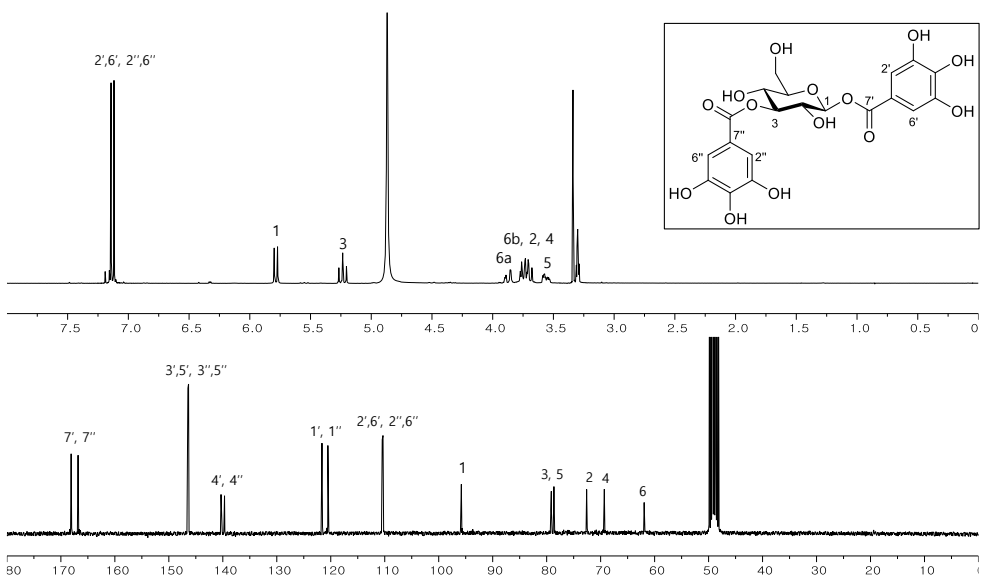


Figure 33. ^1H and ^{13}C NMR spectra of compound **14**

Compound **15**, a white amorphous powder, was determined to have a molecular formula of $C_{20}H_{20}O_{14}$ by its negative HRESIMS peak at *m*/*z* 483.0782 $[M - H]^-$

(calcd for $C_{20}H_{19}O_{14}$, 483.0775). On comparing its ^{1}H NMR data with that of **14**, **15** also had two galloyl groups [δ_H, 7.16, 7.12, (each 2H, s), Gal-H-2′, 6′, Gal-H-2″, 6″] and a *β*-glucopyranose moiety [δ_H, 5.73 (1H, *d*, *J* = 7.7), H-1], but chemical shift of H-3 (δ_H 3.56 – 3.49) was found to move to upfield and chemical shift of H-6 (δ_H 4.55 and 4.38) move to downfield (Figure 34). Taken together, compound **15** was determined as 1,6-di-*O*-galloyl-*β*-D-glucose (Foo, 1993).

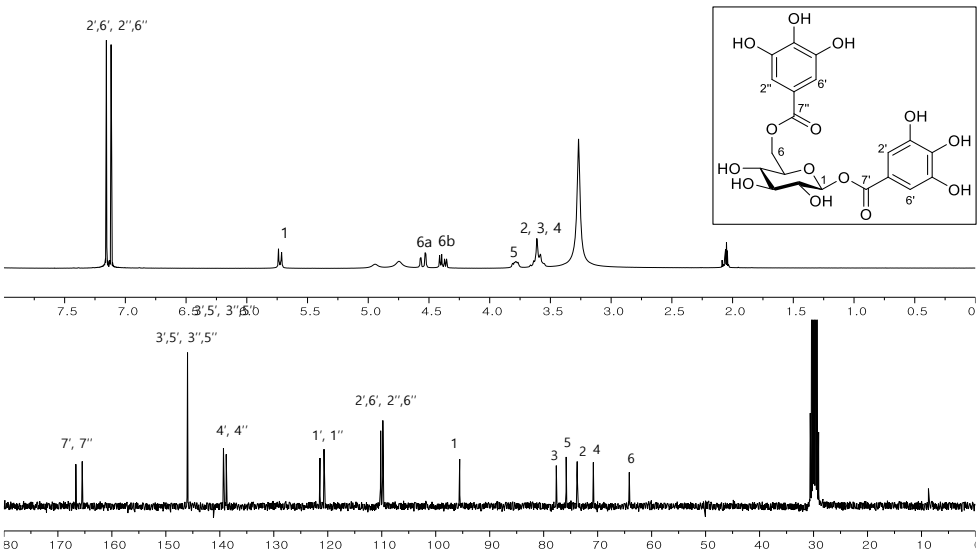


Figure 34. ^{1}H and ^{13}C NMR spectra of compound **15**

3.2. Compounds **16** and **17**

Compound **16** was obtained as light brown amorphous powder. The (-)-HRESIMS of **16** showed a peak at *m*/*z* 613.1219 $[M-H]^-$ (calcd for $C_{29}H_{25}O_{15}$ 613.1193) and yielded a molecular formula of $C_{29}H_{26}O_{15}$. Its ^{1}H NMR spectum revealed the presence of two galloyl groups [δ_H, 7.11, 7.05, (each 2H, s), Gal-H-2′, 6′, Gal-H-2″, 6″] and a *β*-glucopyranose moiety [δ_H, 5.86 (1H, *d*, *J* = 8.3), H-1]. Furthermore, five aromatic protons (δ_H 7.61 – 7.38) and trans olefinic signals [δ_H, 7.72 and 6.53, (each 1H, *d*, *J* = 16.0), Cin-H-7′ and Cin-H-8′] indicated that **16** had a cinnamoyl unit (Figure 35). The H-1 (δ_H 5.86) and H-6 (δ_H 4.60 and 4.37) signals

were correlated to the galloyl carbonyl carbon signals at 167.1 and 167.0 (Gal-C-7′ and Gal-C-7″) in the HMBC spectrum. In addition, the H-2 signal (δ_H 5.18) was correlated to the cinnamoyl carbonyl carbon signal at δ_C 165.6 in the HMBC spectrum. Thus, compound **16** was confirmed as 1,6-di-*O*-galloyl-2-*O*-cinnamoyl-*β*-D-glucose (Kashiwada et al., 1984).

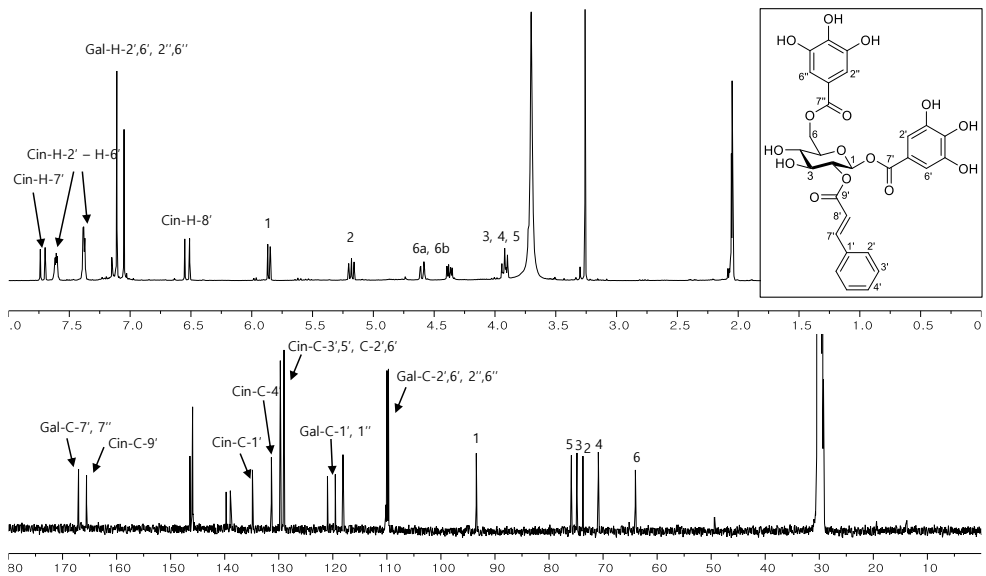


Figure 35. ^{1}H and ^{13}C NMR spectra of compound **16**

Compound **17**, which had the same molecular formula as **16** according to its (-)-HRESIMS [*m*/*z* 613. 1210 $[M - H]^-$ (calcd for $C_{29}H_{25}O_{15}$. 613.1193)], was obtained as brown amorphous powder. Its ^{1}H NMR spectrum and molecular formula suggested that **17** also have two galloyl, one cinnamoyl and one *β*-glucopyranose moieties (Figure 36). The locations of the two galloyl and one cinnamoyl groups were determined to be C-1, C-2 and C-6 positions of glucose, respectively, by analysis of the HMBC spectrum of **17**. From obtained spectral data and above analysis, compound **17** was confirmed as 1,2-di-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (Kashiwada et al., 1984).

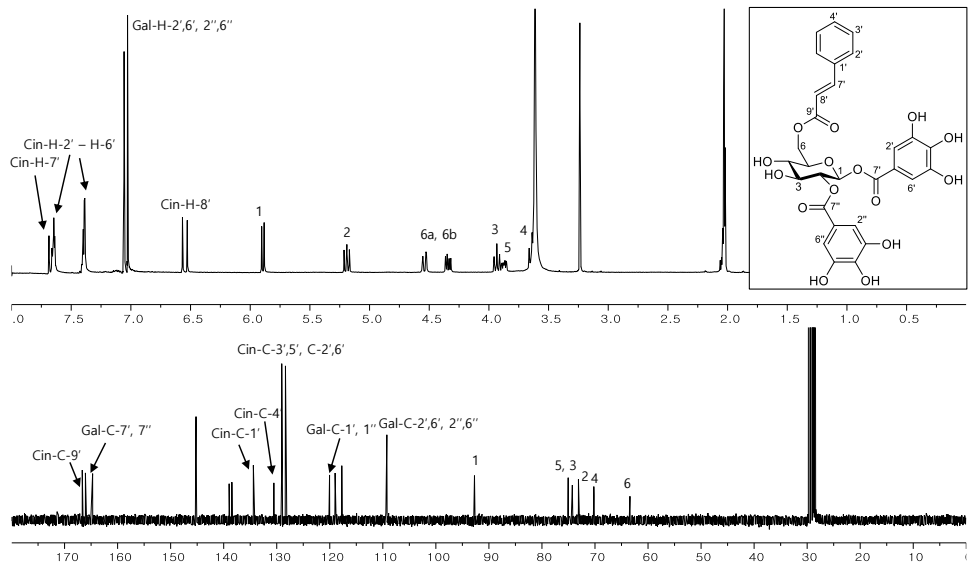


Figure 36. ^{1}H and ^{13}C NMR spectra of compound **17**

3.3. Compound **18**

Compound **18** was obtained as a light brown amorphous powder and had the molecular formula $C_{42}H_{32}O_{27}$ as determined by HRESIMS at m/z 967.1055 [M – H]$^{-}$ (calcd. for $C_{42}H_{31}O_{27}$, 967.1053). The ^{1}H NMR spectrum (Figure 37) revealed signals at δ_H 7.19 (2H, *s*, Gal-H-7′), 7.08 (1H, *s*, HHDP-H-5′) and 6.67 (1H, *s*, HHDP-H-5″) which indicate the existence of one galloyl ester unit and one HHDP unit, respectively. The $^{1}C_4$ conformation of the glucopyranose moiety was confirmed from seven signals with a discrete ^{1}H–^{1}H COSY system and the small coupling constants for the signals of H-1 – H-4 of glucose [δ_H 6.51 (1H, *s*, H-1), 5.90 (1H, *s*, H-3), 5.51 (1H, *s*, H-2), 5.23 (1H, *d*, *J* = 3.5 Hz, H-4)]. In contrast to other neochebuloyl unit containing compounds, the chemical shift of H-2 in **18** moved downfield to δ_H 5.51, which means that C-2 of glucose was acylated. In addition to the galloyl unit, HHDP unit and glucose, the ^{1}H NMR spectrum showed one methylene at δ_H 2.17 (2H, *d*, *J* = 7.7 Hz, Cheb-H-5′) and three methine protons at δ_H 5.13 (1H, *dd*, *J* = 7.2, 0.8 Hz, Cheb-H-3′), 4.97 (1H, *d*, *J* = 7.2 Hz, Cheb-H-2′)

and 3.90 (1H, *t*, *J* = 8.2 Hz, Cheb-H-4′) with a discrete ^{1}H-^{1}H COSY system, one-proton singlet at δ_H 7.51 (1H, *s*, Cheb-H-2″) and one methoxy protons at δ_H 3.54 (3H, *s*). Except for the methoxy protons, 1D- and 2D-NMR spectral data were similar to the chebulagic acid (**3**) which did not contain a 4-*O*-neochebuloyl unit but instead a 2-4-*O*-chebuloyl unit. The H-3 (δ_H 5.90) and H-6 (δ_H 4.78 and 4.41) of glucose were correlated to the HHDP carbonyl carbon signals at δ_C 166.2 (HHDP-C-7′) and 168.3 (HHDP-C-7″) in the HMBC spectrum (Figure 39), respectively, which indicated that the HHDP unit was connected to C-3 and C-6 of the central glucose. Furthermore, H-1 (δ_H 6.51) was correlated to the galloyl carbonyl carbon signals at δ_C 164.9 (Gal-C-7′). The correlation between H-2 (δ_H 5.51) and H-4 (δ_H 5.23) of glucose with the chebuloyl carbonyl carbon at δ_C 165.4 (Cheb-C-7″) and 173.5 (Cheb-C-7′) showed that the chebuloyl unit was connected to C-2 and C-4 of the central glucose. The methoxy proton (δ_H 3.54) was correlated to C-6′ of the chebuloyl unit (δ_C 172.2) in the HMBC spectrum. The (*R*)-configuration of HHDP group was confirmed by CD spectrum of corilagin (**42**), which was isolated from a partial hydrolysate of **18**, according to a strong negative Cotton effect at 237 nm. The absolute configuration of the chebuloyl unit was determined as (2′*S*, 3′*S*, 4′*S*) by comparing $J_{2'3'}$ and $J_{3'4'}$ couplings with those of **18** and by observing the Overhauser effects between H-1 (δ_H 6.51) and Cheb-H-3′ (δ_H 5.13) in its ROESY spectrum (Figure 40) (Pfundstein et al., 2010). The presence of the D-glucose was confirmed by HPLC analysis with Tanaka's method on the acid hydrolysate of **18** (Tanaka et al., 2007). Thus, compound **17** was characterized as methyl chebulagate and first isolated from nature.

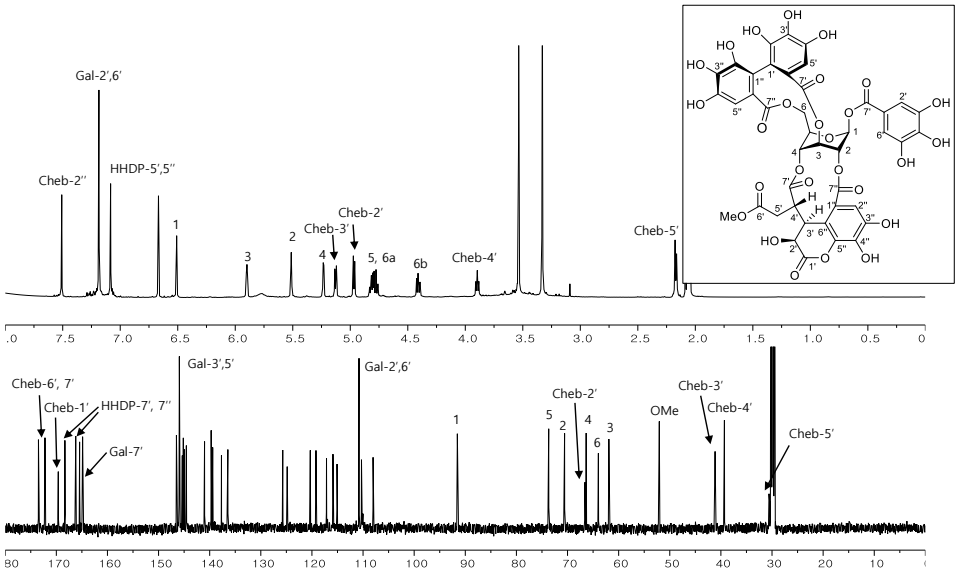


Figure 37. ^{1}H and ^{13}C NMR spectra of compound **18**

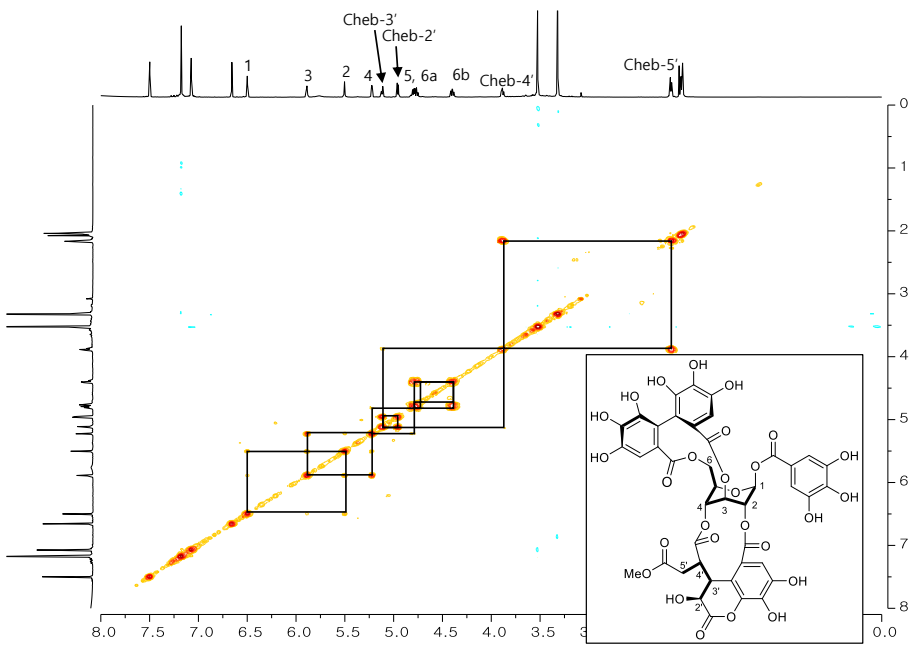


Figure 38. ^{1}H-^{1}H COSY spectrum of compound **18**

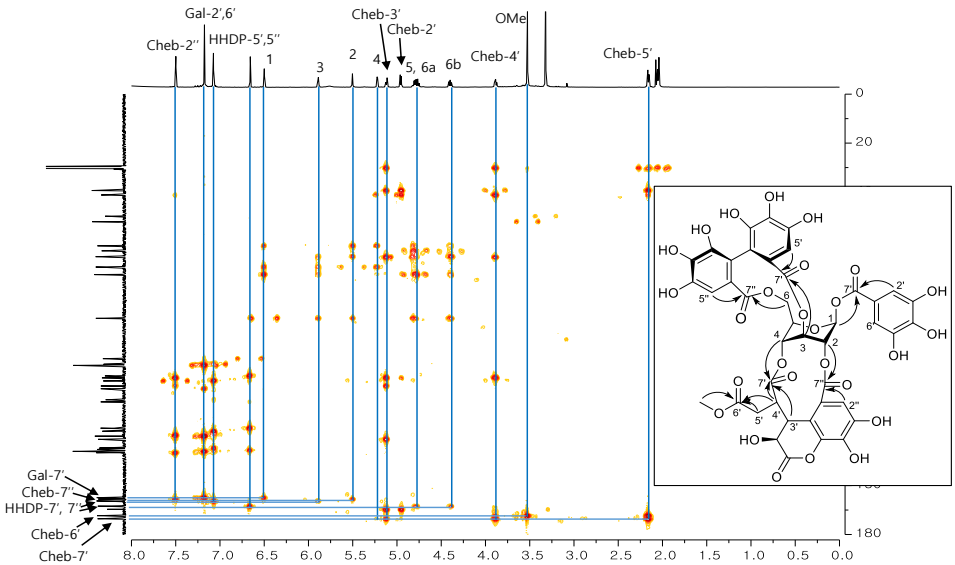


Figure 39. HMBC spectrum of compound **18**

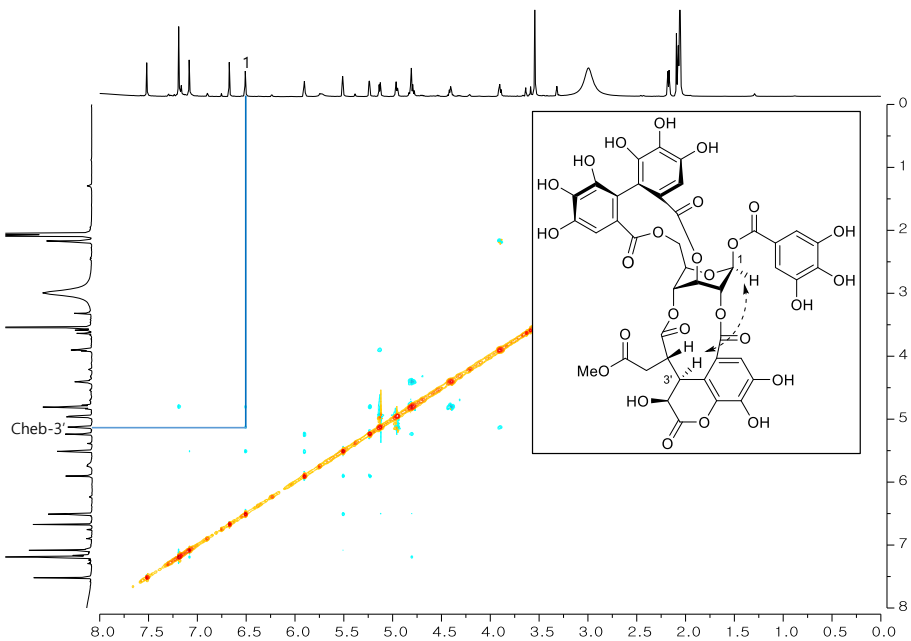


Figure 40. ROESY spectrum of compound **18**

3.4. Compounds **19**-**21**

Compound **19** was obtained as yellowish amorphous powder and determined to have a molecular formula of $C_{27}H_{26}O_{20}$ based on the negative HRESIMS peak at m/z 669.0928 $[M - H]^-$ (calcd for $C_{27}H_{25}O_{20}$, 669.0939). The ^1H NMR spectrum (Figure 41) of **19** showed signals at δ_H 7.19 (2H, *s*, Gal-2′,6′) indicating the existence of a galloyl unit. The large coupling constants (8.1–9.8 Hz) of the H-1 – H-4 signals indicated the β-glucopyranose with $^4C_1$ conformation. In addition to a galloyl unit and glucose, the presence of neochebuloyl moiety was deduced from one methylene at δ_H 3.07 (1H, *dd*, *J* = 12.0, 17.0 Hz, Neocheb-H-5′a) and 2.37 (1H, *dd*, *J* = 3.0, 17.0 Hz, Neocheb-H-5′b) and three methine protons at δ_H 5.70 (1H, *d*, *J* = 1.5 Hz, Neocheb-H-2′), 3.88 (1H, *dd*, *J* = 1.5, 10.5 Hz, Neocheb-H-3′) and 3.14 (1H, *ddd*, *J* = 3.0, 10.5, 12.0 Hz, Neocheb-H-4′) with a discrete ^1H–^1H COSY system, and one-proton singlet at δ_H 7.13 (1H, *s*, Neocheb-H-2″). The H-1 (δ_H 5.79) and H-4 (δ_H 5.07) signals of glucose were correlated to the galloyl carbonyl carbon at δ_C 165.2 (Gal-C-7′) and neochebuloyl carbonyl carbon at δ_C 166.3 (Neocheb-C-7′) in the HMBC spectrum, respectively. On the basis of above information, compound **19** was determined as phyllanemblinin E (Zhang et al., 2001).

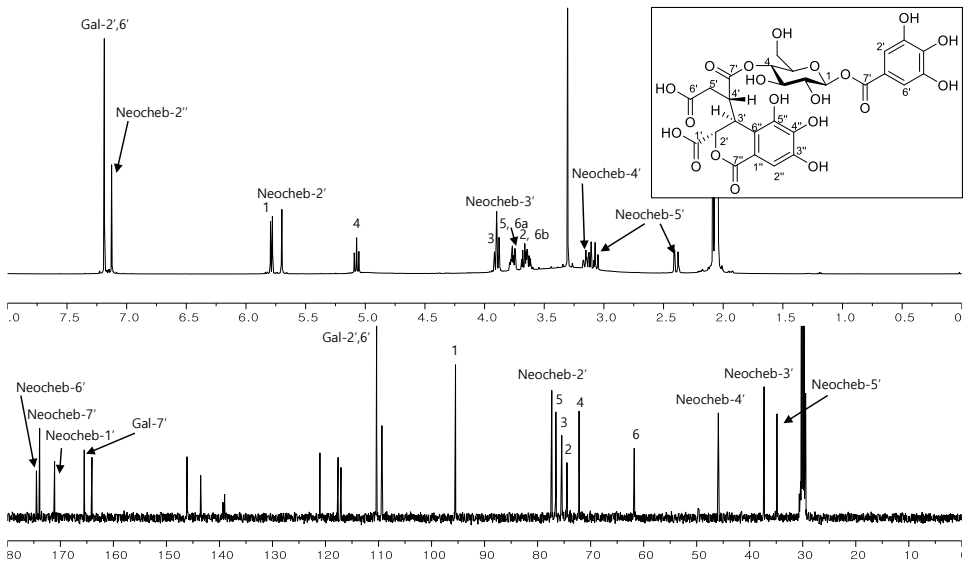


Figure 41. ^{1}H and ^{13}C NMR spectra of compound **19**

Compound **20** was obtained was obtained as a white amorphous powder and had the molecular formula $C_{28}H_{28}O_{20}$ determined by the HRESIMS at m/z 683.1098 $[M - H]^-$ (calcd. for $C_{28}H_{27}O_{20}$, 683.1096). The ^{1}H NMR spectrum (Figure 42) of **20** showed signals at δ_H 7.16 (s, 2H, Gal-2′,6′) indicating the existence of a galloyl unit. The large coupling constants (8.2–9.7 Hz) of the H-1 – H-4 signals of glucose indicated the β-glucopyranose with $^{4}C_1$ conformation [δ_H 5.72 (1H, *d*, *J* = 8.1 Hz, H-1), 4.99 (1H, *t*, *J* = 9.7 Hz, H-4), 3.85 (1H, *t*, *J* = 9.3 Hz, H-3), 3.64 (1H, *t*, *J* = 8.6 Hz, H-2)]. In addition to the galloyl unit and glucose, the ^{1}H NMR spectrum showed one methylene at δ_H 2.97 (1H, dd, *J*=17.2 and 11.6 Hz, H-5′a) and 2.28(1H, dd, *J*=17.2 and 3.2 Hz, H-5′b) and three methine protons at δ_H 5.68 (1H, d, *J*=0.9 Hz, H-2′), 3.78 (1H, dd, *J*=10.5, 0.9 Hz, H-3′) and 3.07 (1H, td, *J*=11.3, 3.2 Hz, H-4′) with a discrete ^{1}H–^{1}H COSY system, one-proton singlet at δ_H 7.09 (1H, s, H-2″) and one methoxy protons at δ_H 3.53 (3H, *s*). Except for the methoxy protons, the coupling constants and chemical shifts of these signals were almost the same as the compound **20**. In the HMBC spectrum, the methoxy protons

were correlated to C-1′ of the neochebuloyl unit (δ_C 170.6). The HMBC analysis (Figure 44) also confirmed that the galloyl and neochebuloyl units were connected to C-1 and C-4 of glucose, respectively. The absolute configuration of the neochebuloyl unit was determined as (2′*S*, 3′*S*, 4′*S*) according to a positive Cotton effect at 235 nm in CD spectrum of chebulic acid from acid hydrolysate of **20**. The D-configuration of the glucose was confirmed by acid hydrolysis with the same methodology used for **18**. Thus, compound **20** was established as 1′-*O*-methyl neochebulanin. The presence of methyl neochebulanin in *T. chebula* was predicted by Pfundstein et al. using HPLC-ESI-MS, but the spectroscopic data, including the NMR, is reported for the first time in this paper (Pfundstein et al., 2010).

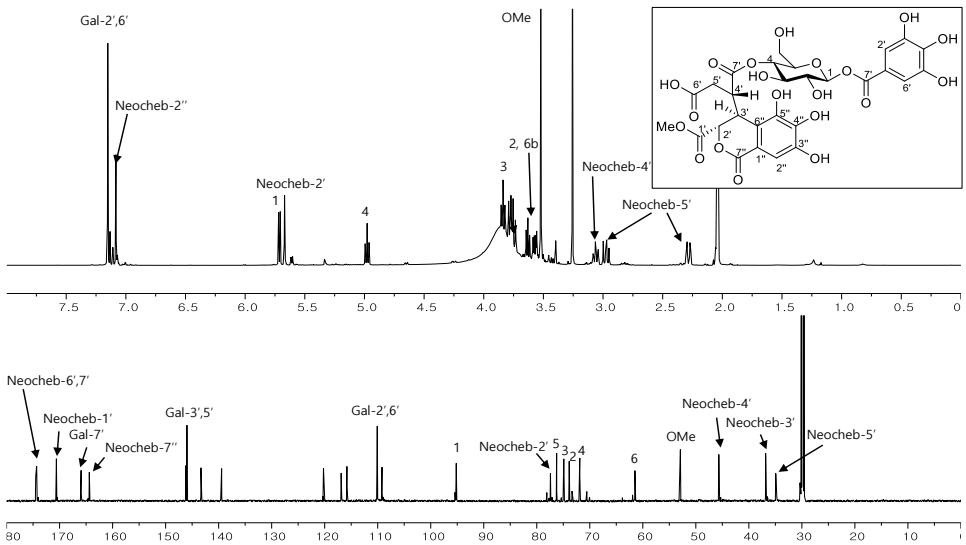


Figure 42. ^{1}H and ^{13}C NMR spectra of compound **20**

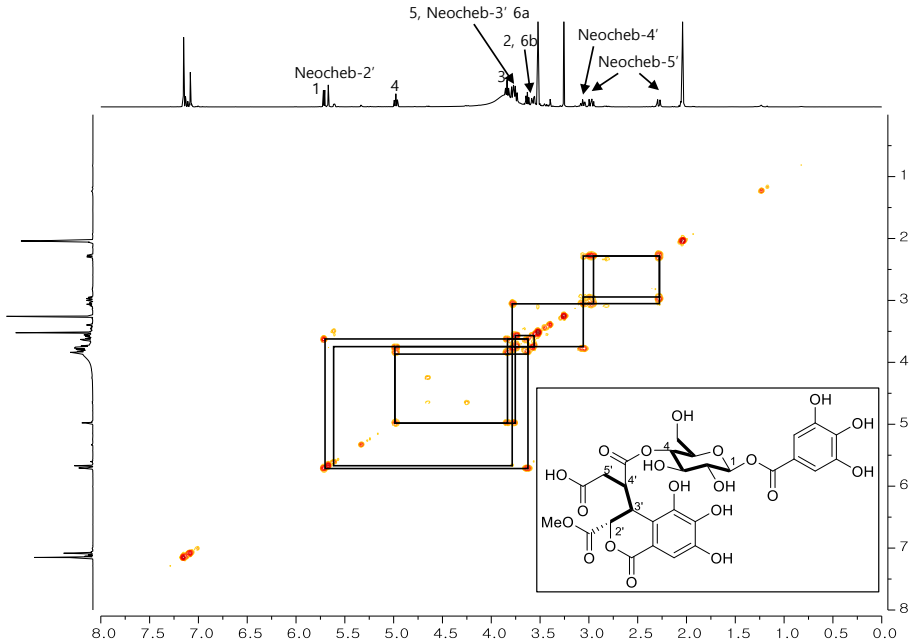


Figure 43. ^{1}H-^{1}H COSY spectrum of compound **20**

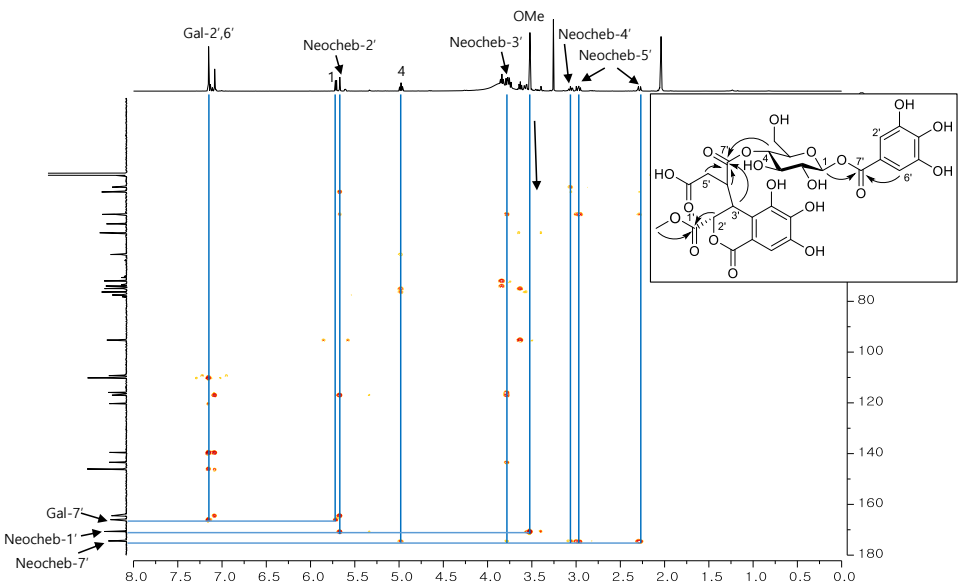


Figure 44. HMBC spectrum of compound **20**

Compound **21** was obtained as a yellowish amorphous powder and had the

molecular formula $C_{43}H_{38}O_{28}$ as determined by the HRESIMS at m/z 1001.1484 $[M - H]^-$ (calcd. for $C_{43}H_{37}O_{28}$, 1001.1472). The ^{1}H-NMR spectrum (Figure 45) of **21** showed signals at δ_H 7.19, 7.17 and 7.16 (each 2H, s, Gal-H-2′,6′, Gal-H-2‴,6‴, Gal-H-2″,6″), which suggested the presence of three galloyl units. The existence of *β*-glucopyranose with $^4C_1$ conformation was apparent from the large coupling constants (8.2–9.8 Hz) of the H-1, H-3 and H-4 signals [δ_H 6.05 (1H, *d*, *J* = 8.2 Hz, H-1), 5.67 (1H, *t*, *J* = 9.5 Hz, H-3), 5.54 (1H, *t*, *J* = 9.8 Hz, H-4)] of central glucose. Compound **21** also had one methylene at δ_H 2.80 (1H, *dd*, *J* = 17.5, 9.1 Hz, Neocheb-H-5′a) and 2.43 (1H, *dd*, *J* = 17.5, 5.2 Hz, Neocheb-H-5′b), three methine protons at δ_H 4.81 (1H, *s*, Neocheb-H-2′), 3.78 (1H, *d*, *J* = 11.2 Hz, Neocheb-H-3′) and 3.20 (1H, *ddd*, *J* = 5.2, 9.1, 11.2 Hz, Neocheb-H-4′) with a discrete ^{1}H–^{1}H COSY system, one-proton singlet at δ_H 7.06 (1H, *s*, Neocheb-H-2″) and two methoxy protons at δ_H 3.34 and 3.32 (each 3H, *s*). Based on these results, the coupling constants and chemical shifts of **21** were similar to neochebulinic acid (Qi et al., 2013) and 1′-*O*-methyl neochebulinate (**4**). The H-1 (δ_H 6.05), H-3 (δ_H 5.67), H-6a (δ_H 4.56) and H-6b (δ_H 4.45) signals of glucose were correlated to the galloyl carbonyl carbon signals at δ_C 165.2 (Gal-C-7′), 166.3 (Gal-C-7″), and 166.5 (Gal-C-7″) in the HMBC spectrum (Figure 47), respectively. The HMBC correlation of H-4 (δ_H 5.54) with neochebuloyl carbonyl carbon (δ_C 173.5, Neocheb-C-7′) revealed that the 4-OH group of the central glucose was esterified with a neochebuloyl unit. The HMBC spectrum also showed the correlations of two methoxy protons to C-1′ (δ_C 169.7) and C-6′ (δ_C 172.2) of the neochebuloyl unit. The configuration of the neochebuloyl unit was determined as (2′*S*, 3′*S*, 4′*S*) according to a positive Cotton effect at 236 nm in the CD spectrum of chebulic acid from acid hydrolysate of **21**. Acid hydrolysis of **10** yielded D-glucose which was determined by the method mentioned above. Thus, compound **10** was characterized as dimethyl neochebulinate and isolated for the first time from nature.

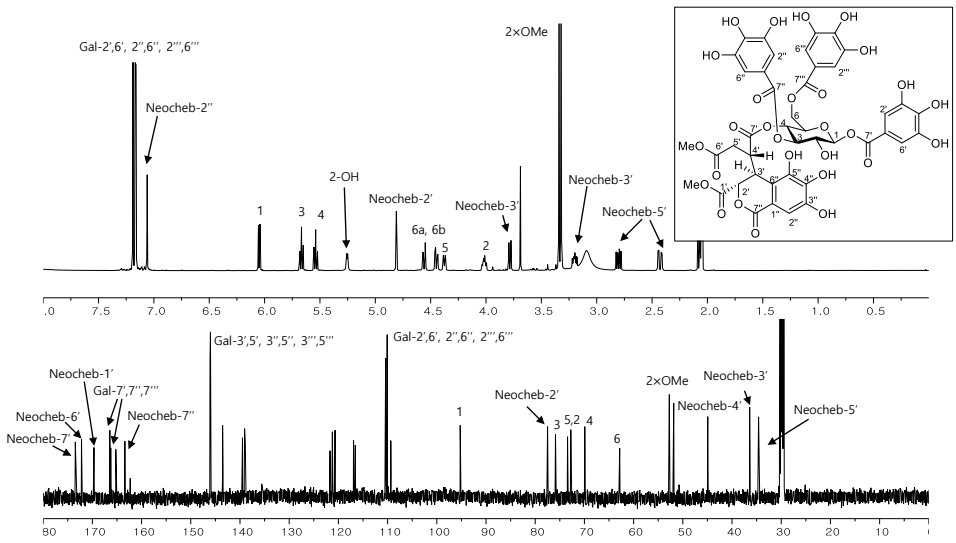


Figure 45. ^{1}H and ^{13}C NMR spectra of compound **21**

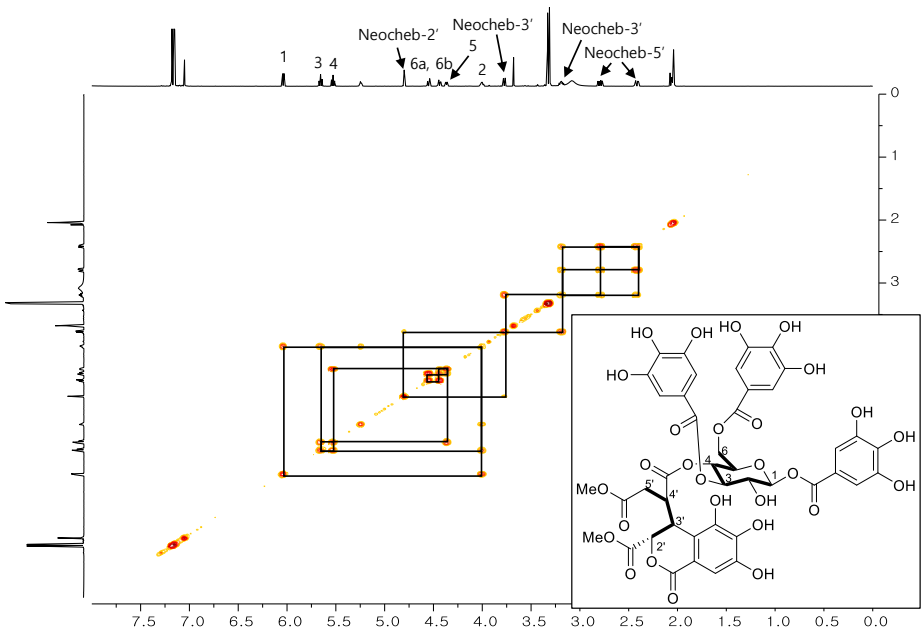


Figure 46. ^{1}H-^{1}H COSY spectrum of compound **21**

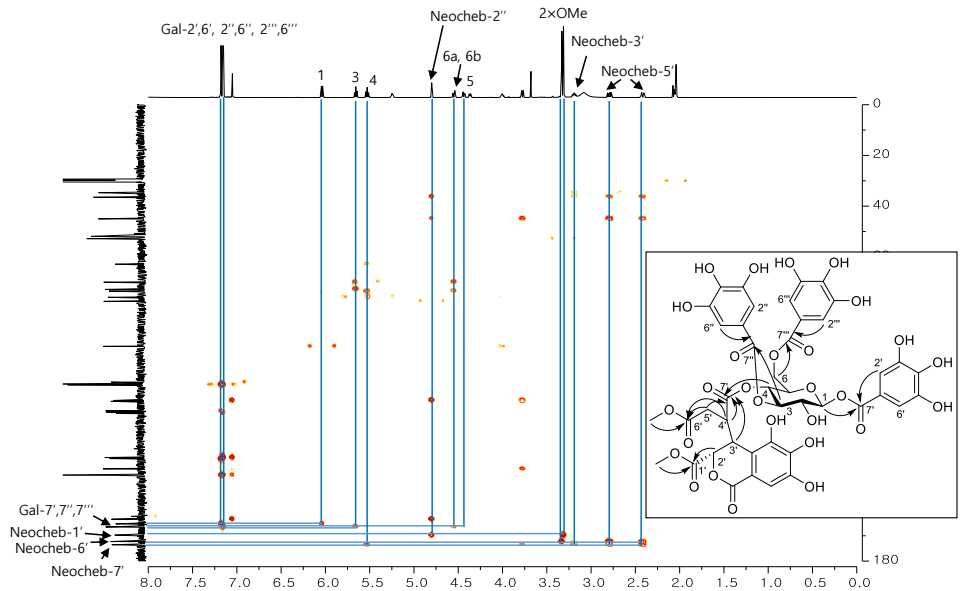


Figure 47. HMBC spectrum of compound **21**

3.5. Compounds **22**-**24**

Compound **22** was obtained as yellowish amorphous powder and with molecular formula $C_{41}H_{32}O_{28}$, indicated by (-)-HRESIMS at *m*/*z*: 971.0989 $[M - H]^-$ (calcd for $C_{41}H_{31}O_{28}$, 971.1002). The ^{1}H NMR spectrum of **22** showed signals at δ_H 7.10 (2H, *s*, Gal-H-2′,6′) indicating the presence of a galloyl ester unit. In compound **22**, compared to **21**, two galloyl units formed a hexahydroxydiphenoyl (HHDP) unit with evidence from the ^{1}H-NMR signals at δ_H 6.83 and 6.70 (each 1H, *s*, HHDP-H-5′, 5″). The *β*-glucopyranose with $^1C_4$ conformation was deduced from seven signals with a discrete ^{1}H-^{1}H COSY system and the coupling constants (3.0–3.2 Hz) of the H-1 and H-4 signals [δ_H 6.28 (1H, *d*, *J* = 3.0 Hz, H-1) and 5.70 (1H, *d*, *J* = 3.2 Hz, H-4)] of glucose and carbon signal of C-1 at δ_C 94.7. The ^{1}H–^{1}H COSY spectrum also showed consecutive correlations from H-2′ (δ_H 5.50) to H-5′ (δ_H 2.86 and 2.46) through H-3′ (δ_H 3.99) and H-4′ (δ_H 3.28) in the neochebuloyl unit. HMBC correlations were observed between carbonyl carbon signals at δ_C 168.3

(HHDP-C-7′) and 170.2 (HHDP-C-7″) and the protons at δ_H 6.75 (HHDP-H-5″) and 6.74 (HHDP-H-5″) as well as between the carbonyl carbon signals at δ_C 166.7 (Gal-C-7′) and the galloyl H-2′ and 6′ proton resonances at δ_H 7.07. The H-3 (δ_H 4.90) and H-6 (δ_H 4.85 and 4.27) signals of glucose were correlated to the HHDP carbonyl carbon signals at δ_C 168.3 and 170.2 in the HMBC spectrum, respectively, and indicated that the HHDP unit was connected to C-3 and C-6 of the central glucose. Similarly, it was confirmed that the galloyl unit was connected to C-1 of the glucose. The correlation between H-4 (δ_H 5.70) of glucose with neochebuloyl carbonyl carbon (C-7′, δ_C 173.6) showed that the neochebuloyl unit was connected to C-4 of the central glucose. With above observed spectroscopic data, compound **22** was confirmed as neochebulagic acid (Saijo et al., 1989).

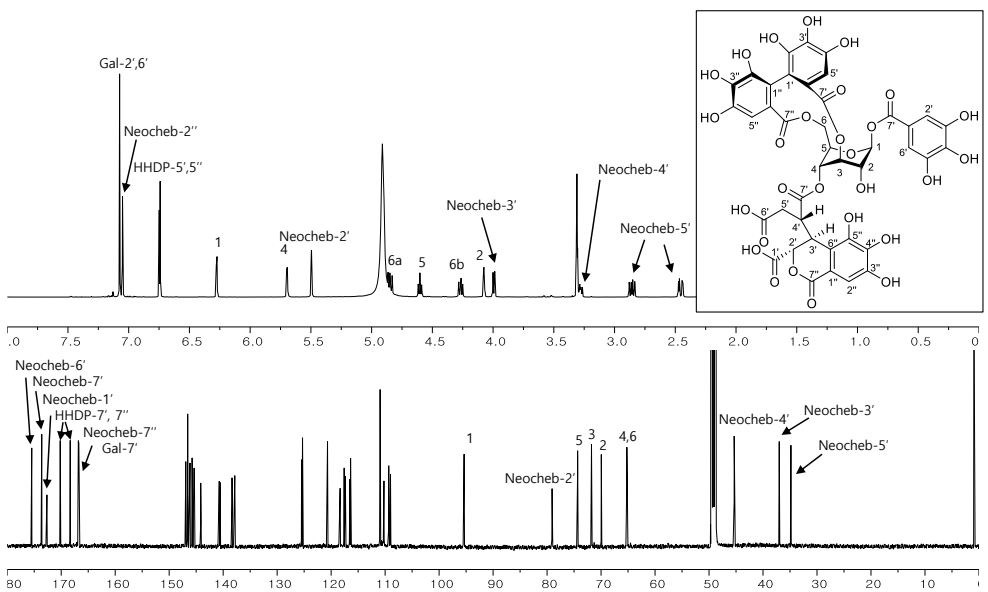


Figure 48. ^{1}H and ^{13}C NMR spectra of compound **22**

Compound **23** was obtained as yellowish amorphous powder and had the molecular formula $C_{43}H_{36}O_{28}$ as determined by HRESIMS at m/z 999.1298 [M – H]$^-$ (calcd. for $C_{43}H_{35}O_{28}$, 999.1315). Except for two methoxy moiety (δ_H 3.55/δ_C

53.0 and δ_H 3.52/δ_C 52.1), the 1D and 2D NMR spectral data of **23** were similar to those of compound **22**. The signal by the methoxy moieties was correlated to C-1′ (δ_C 170.6) and C-6′ (δ_C 172.6) of the neochebuloyl unit in the HMBC spectrum (Figure 51). The absolute configuration of 3,6,-*O*-HHDP group was assigned to be (*R*)-configuration by CD spectrum of corilagin (**42**), which was isolated from a partial hydrolysate of **3**, due to a strong negative Cotton effect at 237 nm (Okuda et al., 1982). To confirm the absolute configuration of the neochebuloyl unit, compound **3** underwent acid hydrolysis and yielded chebulic acid (**34**). The absolute configuration of chebulic acid was determined as (2′*S*, 3′*S*, 4′*S*) according to a positive Cotton effect at 235 nm in the CD spectrum (Yoshida et al., 1982). Acid hydrolysate of **23** affforded D-glucose which was determined by the method mentioned above (Tanaka et al., 2007). Thus, compound **23** was characterized as Dimethyl neochebulagate, which has been newly discovered from nature.

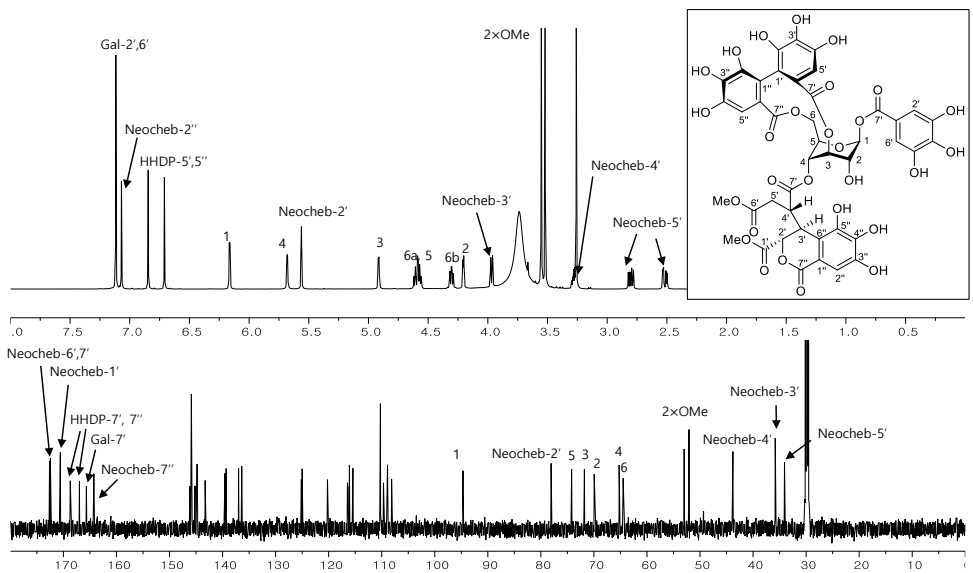


Figure 49. ^{1}H and ^{13}C NMR spectra of compound **23**

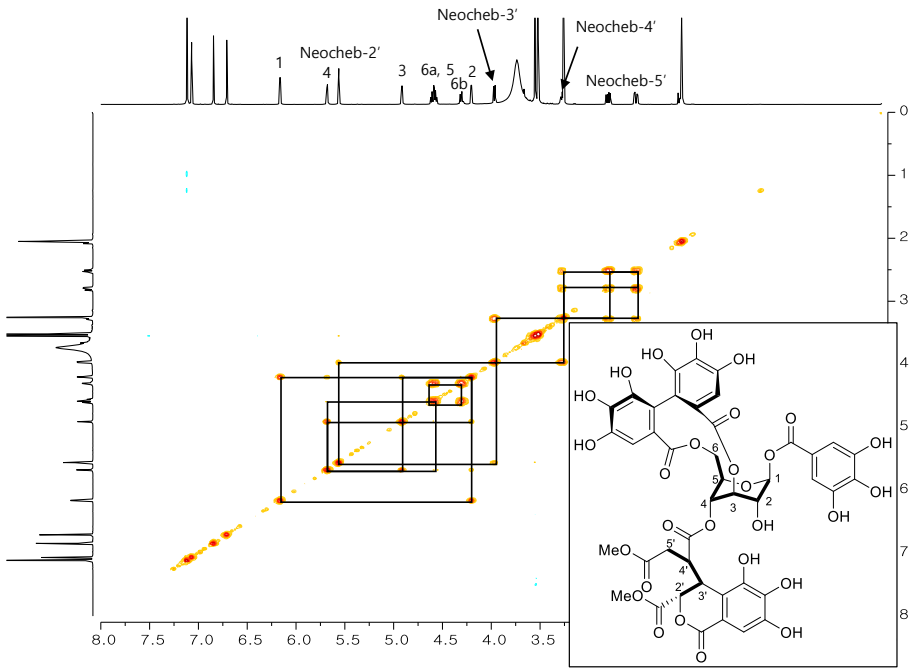


Figure 50. ^{1}H-^{1}H COSY spectrum of compound **23**

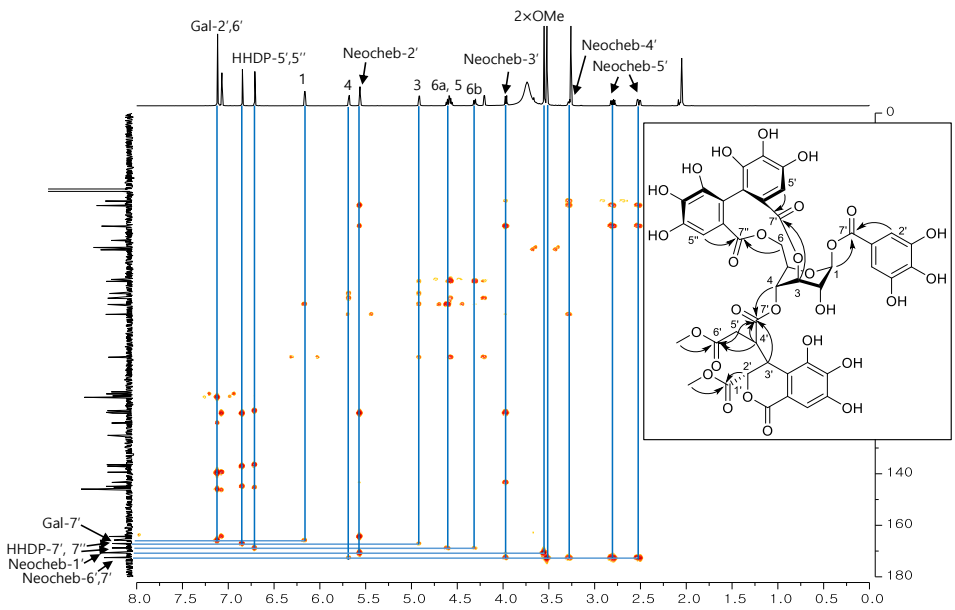


Figure 51. HMBC spectrum of compound **23**

Compound **24** was obtained as a yellowish amorphous powder and had the molecular formula $C_{43}H_{36}O_{28}$ as determined by HRESIMS at m/z 999.1315 $[M - H]^-$ (calcd. for $C_{43}H_{35}O_{28}$, 999.1315). It was similar to the structures of **24** according to the 1D- and 2D-NMR data. Significant differences between **24** and **23** were the shielding of H-2′ (δ_H 5.38, Δ-0.18 ppm) and deshielding of H-3′ (δ_H 4.20, Δ0.23 ppm) and H-4′ (δ_H 3.55, Δ0.27 ppm) of the neochebuloyl unit. Additionally, the coupling constant between H-3′ and H-4′ ($J_{3'4'}$ = 4.2 Hz) of the neochebuloyl unit was smaller than those of **23** ($J_{3'4'}$ = 9.4 Hz) and all the other neochebuloyl moiety containing compounds which have the same configurations (Figure 53). Furthermore, the absence of ROESY cross peaks between H-2′ (δ_H 5.38) and H-4′ (δ_H 3.55) of neochebuloyl moiety provided further evidence for the stereochemistry of C-4′ of neochebuloyl moiety (Figure 56). These spectroscopic differences between **24** and **23** were analogous to those between chebulic acid (**34**) and neochebulic acid (**37**). Neochebulic acid, which was first reported by Ding et al., has the (2′*S*, 3′*S*, 4′*R*) configurations (Ding et al., 2000). Furthermore, 4′-*epi*-neochebulagic acid, which was synthesized by hydrogenation of repandusinic acid A, has similar NMR spectroscopic characteristics to **24** (Saijo et al., 1989). The (*R*)-configuration of HHDP group was confirmed by CD spectrum of corilagin (**42**), which was isolated from a partial hydrolysate of **24**. To confirm the absolute configuration of the 4′-*epi*-neochebuloyl unit, compound **24** underwent acid hydrolysis and gave neochebulic acid which was identified by HPLC analysis and comparison of CD spectra (positive Cotton effect at 213 nm) with an authentic sample (**37**). Acid hydrolysate of **24** also affforded D-glucose which was determined by the method mentioned above. Thus, compound **24** was characterized as dimethyl 4′-*epi*-neochebulagate and isolated first time from nature.

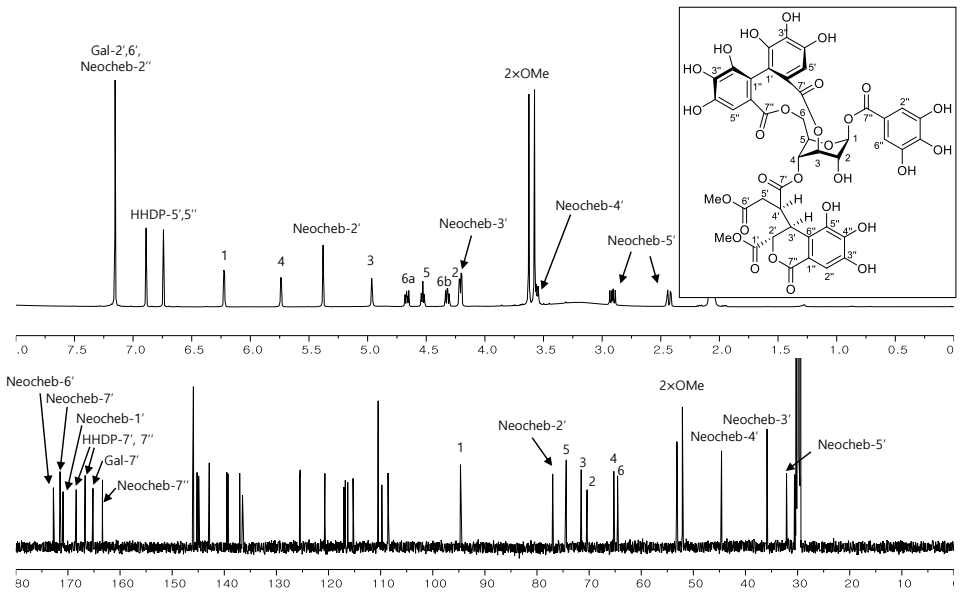


Figure 52. ^{1}H and ^{13}C NMR spectra of compound **24**

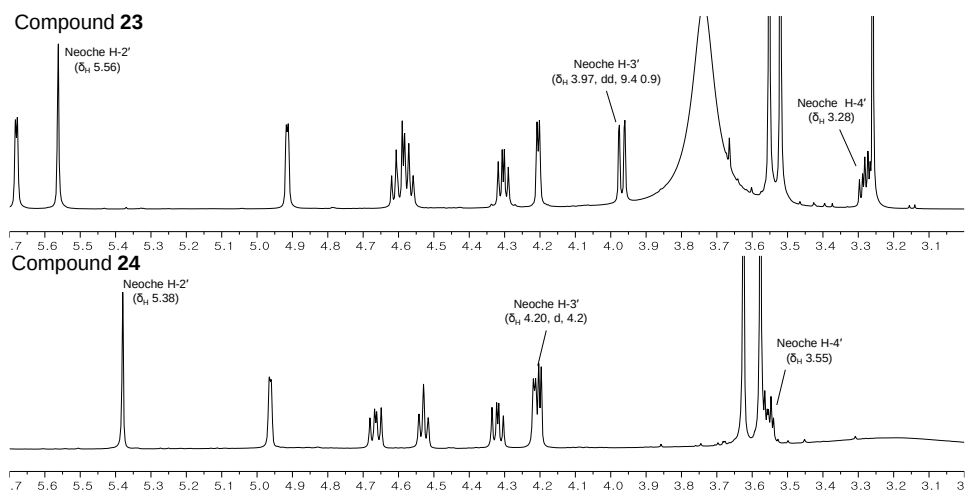


Figure 53. Comparison of expanded ^{1}H NMR spectra between **23** and **24**

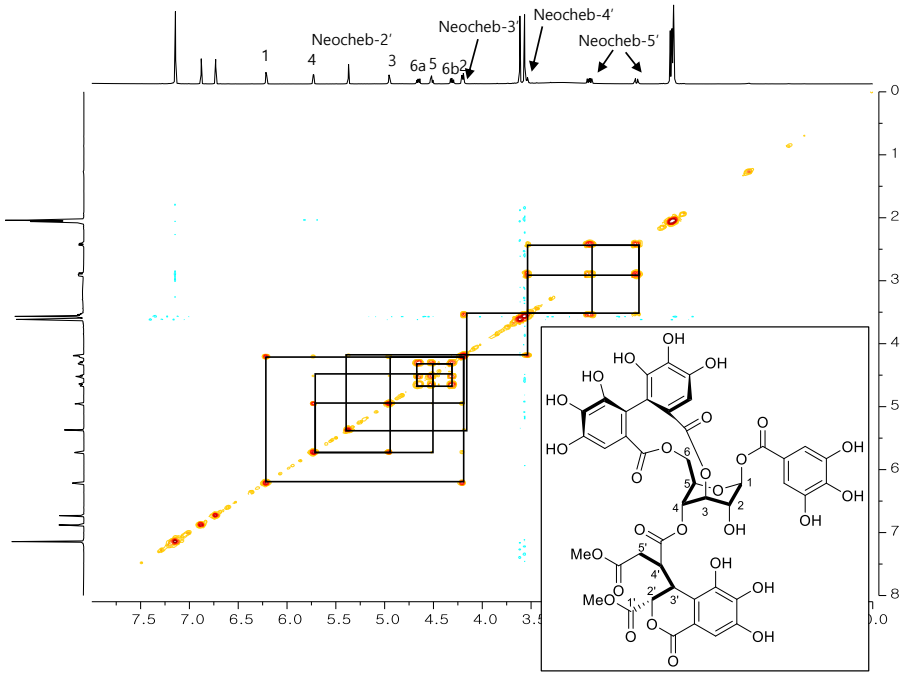


Figure 54. ^{1}H-^{1}H COSY spectrum of compound **24**

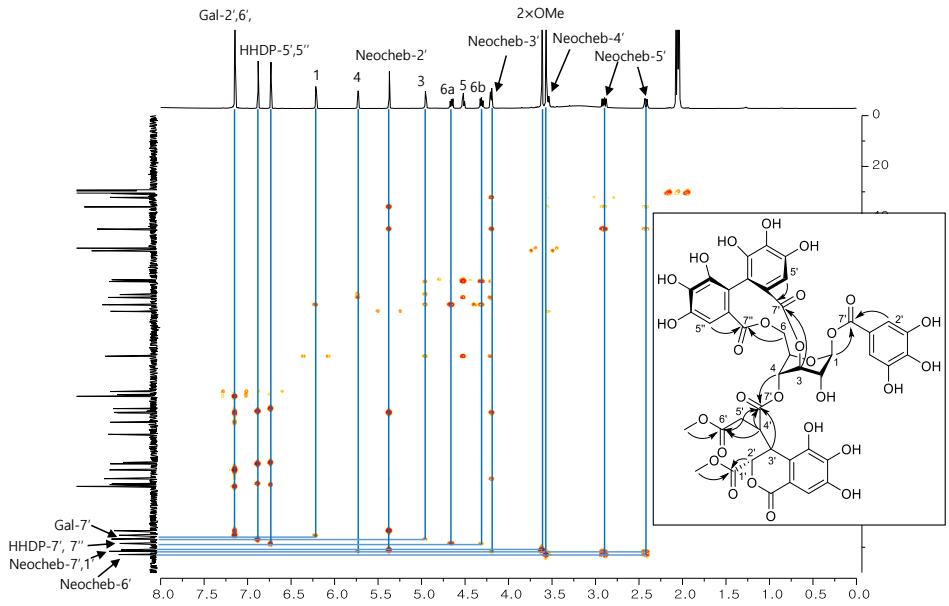


Figure 55. HMBC spectrum of compound **24**

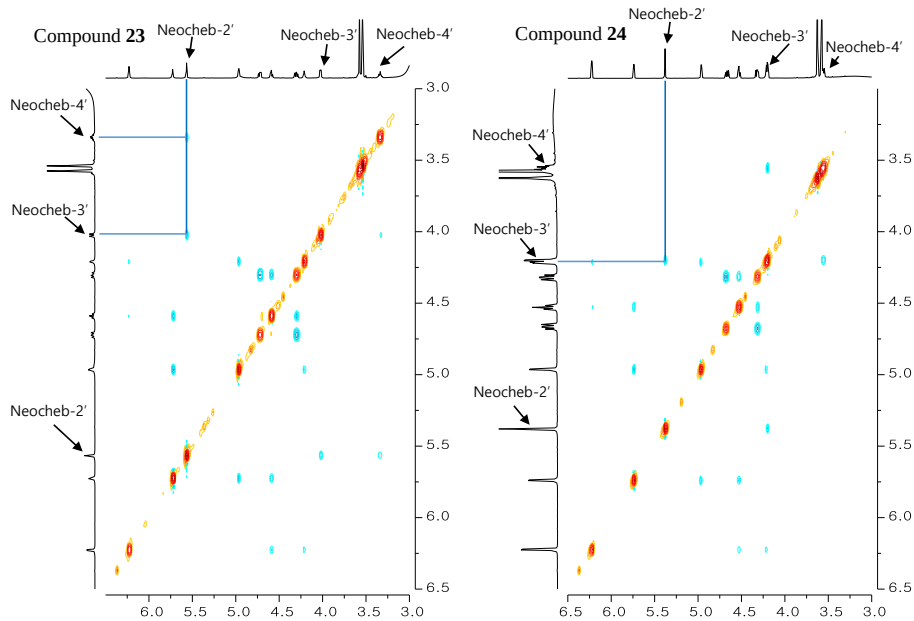


Figure 56. Comparison of expanded ROESY spectra between **23** and **24**

3.6. Compound **25**

Compound **25** was obtained as white amorphous powder and gave a molecular formula of $C_{27}H_{24}O_{19}$ as assigned by the negative HRESIMS peak at *m/z* 651.0825 $[M - H]^-$ (calcd for $C_{27}H_{23}O_{19}$, 651.0834). Its ^1H NMR spectrum (Figure 57) revealed signals for a galloyl [δ_H, 7.13, (2H, s), Gal-H-2′, 6′] and a chebuloyl moiety [δ_H 7.45 (1H, s, Cheb-H-2″), 5.10 (1H, *dd*, *J* = 1.5, 7.2 Hz, Cheb-H-3′), 4.78 (1H, *m*, Cheb-H-2′), 3.81 (1H, *ddd*, *J* = 1.0, 4.1, 5.5 Hz, Cheb-H-4′), 2.15 (2H, *m*, Cheb-H-5′)]. In addition, the coupling constant of anomeric proton signal at δ_H 6.37 (1H, *d*, *J* = 2.8 Hz, H-1) and the presence of anomeric carbon signal at δ_C 92.9 suggested that glucose moiety was β-orientation with $^1C_4$ conformation. Thus, compound **25** was confirmed as chebulanin by comparison with literatures (Pfundstein et al., 2010).

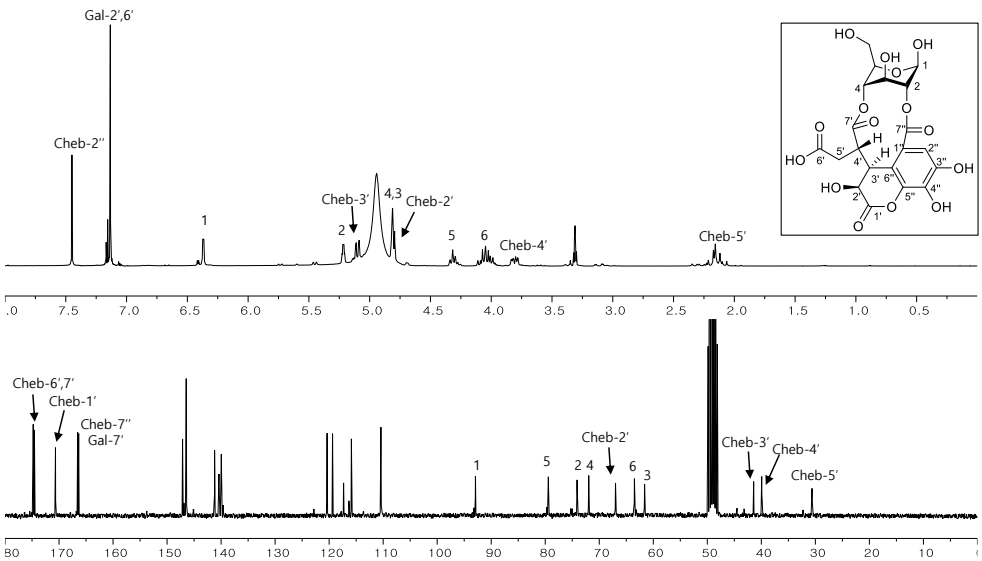


Figure 57. ^{1}H and ^{13}C NMR spectra of compound **25**

3.7. Compound **26**

Compound **26** was obtained as brown amorphous powder and $[\alpha]_D^{20}$= -68.1 (c 0.10, MeOH). The (-)-HRESIMS exhibited quasi-molecular ion peak at *m/z*: 447.0561 $[M - H]^-$ (calcd for $C_{20}H_{15}O_{12}$, 447.0564) indicating $C_{20}H_{16}O_{12}$ as its molecular formula. The presence of ellagic acid moiety was confirmed by its UV spectrum at 255 and 360 nm and two aromatic singlets at δ_H 7.70 and 7.43 (each 1H, H-5 and H-5′). The presence of *α*-rhamnopyranose with $^{1}C_4$ conformation was deduced from five oxygenated methines [δ_H 5.42 (1H, *s*, H-1″), 3.96 (1H, *s*, H-2″), 3.80 (1H, *dd*, *J* = 3.0 9.3 Hz, H-3″), 3.50 (1H, *m*, H-5″), 3.29 (1H, *t*, *J* = 9.4 Hz, H-4″)] along with a doublet methyl proton at δ_H 1.10 (3H, *d*, *J* = 6.2 Hz, H-6″) and the very small vicinal coupling constants of the H-1″ and H-2″ signals (Figure 58). Rhamnosylation at C-4 of ellagic acid was confirmed by the HMBC correlation between the signals of H-1″ (δ_H 5.42) and C-4 (δ_C 146.4). Thus, compound **26** was identified as eschweilenol C (Yang et al., 1998).

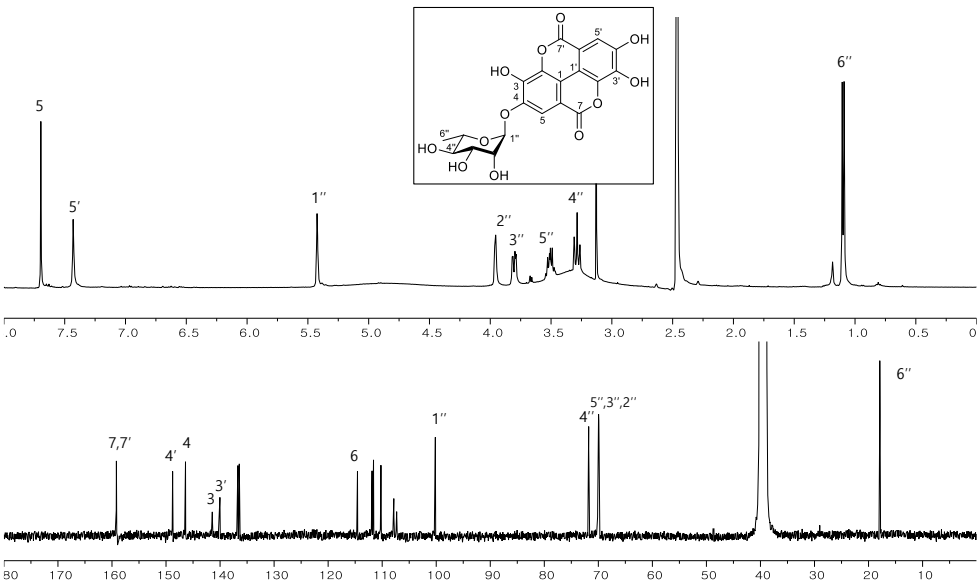


Figure 58. ^{1}H and ^{13}C NMR spectra of compound **26**

3.8. Compounds **27** and **28**

Compound **27** was obtained as white amorphous powder. Its molecular formula, $C_7H_6O_5$, was confirmed from the (-)-HRESIMS data (*m*/*z* 169.0131 $[M - H]^-$, calcd for $C_7H_5O_5$, 169.0137). The ^{1}H and ^{13}C NMR spectra (Figure 59) of **27** displayed only one aromatic proton at δ_H 6.91 (2H, *s*, H-2, 6) and five carbons at δ_C 167.5 (C-7), 145.4 (C-3, 5), 138.0 (C-4), 120.4 (C-1), 108.7 (C-2, 6), respectively. Based on above information, compound **27** was considered to be symmetrical. Thus, compound **27** was confirmed as gallic acid by comparison with literatures (Zhang et al., 2012).

Compound **28** was obtained as yellowish amorphous powder. Its molecular formula, $C_8H_8O_5$, was determined from the (-)-HRESIMS data (*m*/*z* 183.0289 $[M - H]^-$, calcd for $C_8H_7O_5$, 183.0293). ^{1}H and ^{13}C NMR data of **28** showed that its structure was similar to that of **27**, except for a methoxy group [δ_H 3.81 (3H, *s*, OMe), δ_C 52.3 (OMe)]. The locations of the methoxy group was determined to be

C-7 position of gallic acid by analysis of the HMBC spectrum. Thus, compound **28** was identified as methyl gallate (Madikizela et al., 2014).

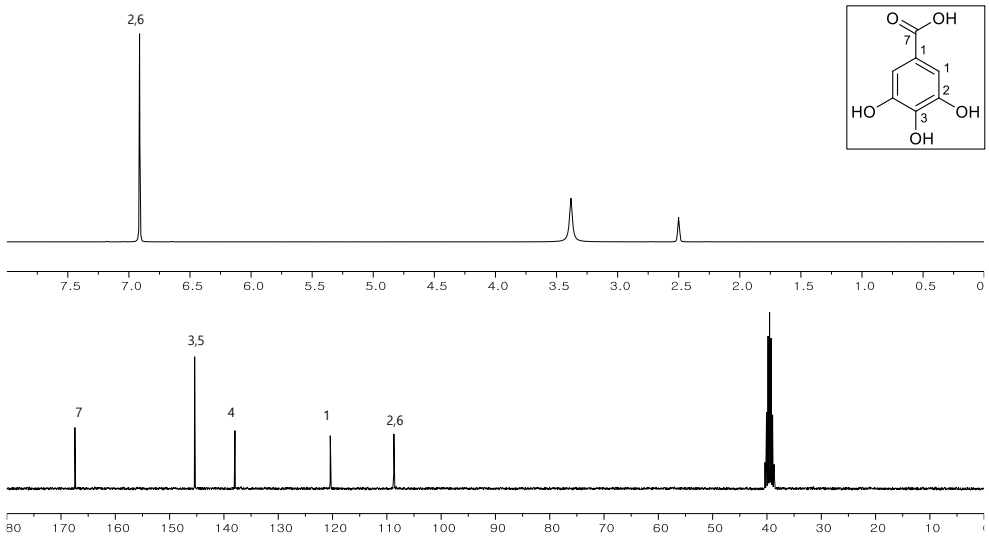


Figure 59. ^{1}H and ^{13}C NMR spectra of compound **27**

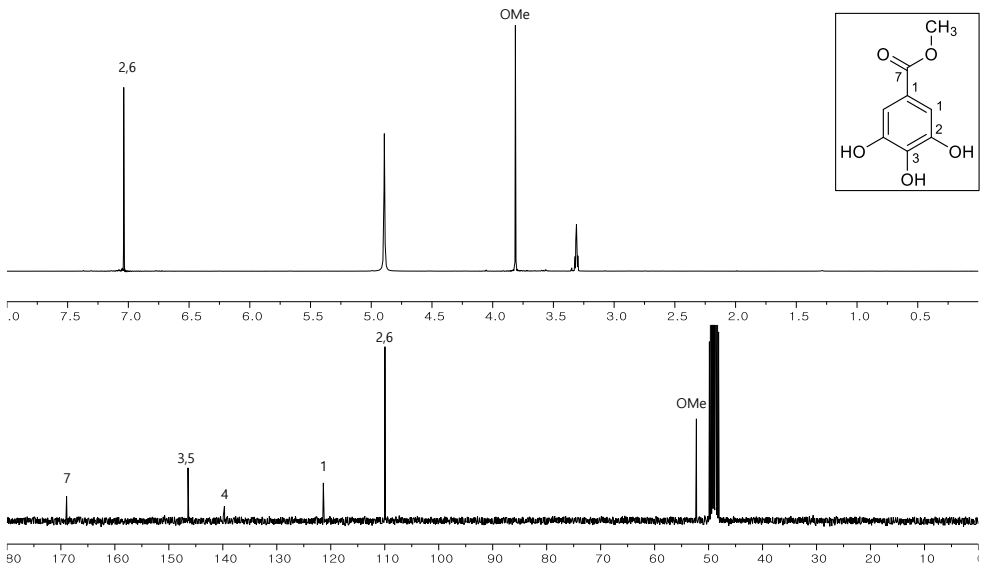


Figure 60. ^{1}H and ^{13}C NMR spectra of compound **28**

3.9. Compounds **29** and **30**

Compound **29** was obtained as yellowish syrup and $[\alpha]_D^{20}$= -224.1 (c 0.10, MeOH). Its molecular formula was identified as $C_{14}H_{14}O_9$ by (-)-HRESIMS data (*m*/*z* 325.0560 [M – H]$^-$, calcd for $C_{14}H_{13}O_9$, 325.0560). Its ^{1}H NMR spectrum showed characteristic signals for a galloyl group at δ_H 7.08 (2H, *s*, H-2′, 6′). In addition, 1D NMR spectra (Figure 61) revealed the presence of a olefinic proton [δ_H 6.84 (1H, *m*, H-6)], three oxygenated methines [δ_H 5.11 (1H, *dd*, *J* = 3.8, 7.3 Hz, H-5), 4.69 (1H, *t*, *J* = 3.7 Hz, H-4), 4.27 (1H, *dt*, *J* = 9.7, 4.8 Hz, H-3)], one methylene [δ_H 2.68, 2.32 (each 1H, m, H-2a, 2b)], and one carboxylic carbon [δ_C 169.1 (C-7)]. The above-mentioned analysis suggested that **29** have a shikimic acid group. The HMBC analysis indicated galloyl group was attached at C-4 of shikimic acid by the correlation from δ_H 4.69 (H-4) to δ_C 167.3 (C-7′). On the basis of above information, compound **29** was determined as 4-*O*-galloyl-(-)-shikimic acid (Ishimaru et al., 1987).

Compound **30**, which had the same molecular formula as **29** according to its (-)-HRESIMS data (*m*/*z* 325.0559 [M – H]$^-$, calcd for $C_{14}H_{13}O_9$, 325.0560) was obtained as yellowish syrup and $[\alpha]_D^{20}$= -165.99 (c 0.10, MeOH). The upfield shifts of H-4 and C-4 (δ_H 4.15 and δ_C 70.1) and the downfield shifts of H-5 and C-5 (δ_H 5.72 and δ_C 70.4) suggested that a galloyl group was located at C-5 in **30** (Figure 62). Thus, compound **30** was elucidated as 5-*O*-galloyl-(-)-shikimic acid (Ishimaru et al., 1987).

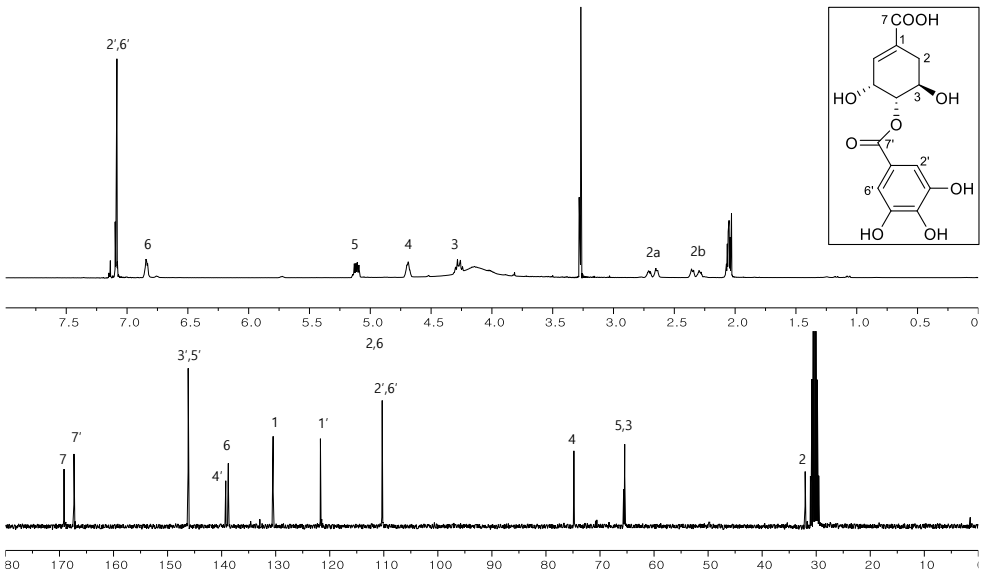


Figure 61. ^{1}H and ^{13}C NMR spectra of compound **29**

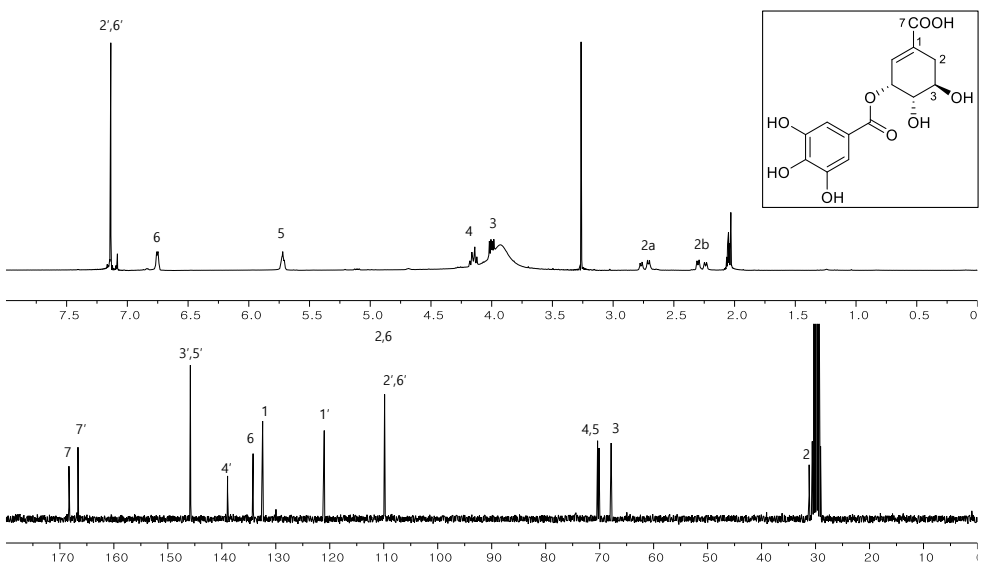


Figure 62. ^{1}H and ^{13}C NMR spectra of compound **30**

3.10. Compounds **31**-**33**

Compound **31** was obtained as yellowish amorphous powder and equilibrium mixture in a ratio of 5:4. Its molecular formula was identified as $C_{13}H_{16}O_{10}$ by (-)-

HRESIMS at *m/z* 331.0649 $[M - H]^-$ (calcd for $C_{13}H_{15}O_{10}$ 331.0665). The ^1H NMR signals at δ_H 5.10 (1H, d, J = 3.7 Hz, *α*-H-1) and 4.54 (0.8H, d, J = 7.8 Hz, *β*-H-1) and corresponding carbon signal at δ_C 93.2 (*α*-C-1) and 97.7 (*β*-H-1) indicated that compound **31** existed anomeric mixtures in solution, namely C-1 was not substituted (Figure 63). The ^1H NMR spectrum also showed one galloyl group [δ_H 7.10 (3.6H, *s, α, β*-H-2′,6′)]. The location of a galloyl moiety was determined to be C-6 position by strongly downfielded ^1H NMR signals of H-6 [δ_H 4.49, 4.32 (*α*-H-6), 4.45, 4.27 (*β*-H-6)]. Thus, compound **31** was identified as 6-*O*-galloyl-D-glucose by comparison with literature (Hsu et al., 1994).

Compound **32** was obtained as yellowish amorphous powder and equilibrium mixture in a ratio of 5:3. Its molecular formula $C_{20}H_{20}O_{14}$, indicated by (-)-HRESIMS (*m/z*: 483.0777 $[M - H]^-$, calcd for $C_{20}H_{19}O_{14}$ 483.0775). Its ^1H and ^{13}C NMR spectra (Figure 64) suggested that compound **32** was also anomeric mitures. In addition, two aromatic singlets at δ_H 7.13 and 7.12 (each 3.2 H, *s, α ,β*-Gal-H-2′,6′, 2″,6″) on its ^1H NMR spectrum evidenced the presence of one additional galloyl unit in compound **32** compared to **31**. The attachment of the two galloyl units was assigned to be the C-3 and C-6 position of central glucose by downfielded signals of H-3 [δ_H 5.39 (*α*-H-3) and 5.16 (*β*-H-3)] and H-6 [δ_H 4.52, 4.41 (*α*-H-6), 4.51, 4.38 (*β*-H-6)] and HMBC correlation between H-3/C-7″ and H-6/C-7′. Therefore, compound **32** was confirmed as 3,6-di-*O*-galloyl-D-glucose.

Compound **33** was obtained as yellowish amorphous powder with molecular formula $C_{27}H_{24}O_{18}$, indicated by (-)-HRESIMS (*m/z*: 635.0889 $[M - H]^-$, calcd for $C_{27}H_{23}O_{18}$ 635.0884) and equilibrium mixture in a ratio of 2:1. Based on the ^1H and ^{13}C NMR spectra (Figure 65), compound **33** was determined to be anomeric mixtures. Its ^1H NMR spectrum indicated that **33** had three galloyl moieties [δ_H 7.14 (2H, *s, α*-Gal-H-2″,6″), 7.13 (1H, *s, β*-Gal-H-2″,6″), 7.02 (3H, *s, α,β*-Gal-H-

2′,6′), 7.01 (2H, *s*, *α*-Gal-H-2‴,6‴), 6.99 (1H, *s*, *β*-Gal-H-2‴,6‴)]. The locations of three galloyl moieties were determined to be C-3, C-4, and C-6 by downfielded signals of H-3 [δ_H 5.70 (*α*-H-3) and 5.51 (*β*-H-3)], H-4 [δ_H 5.34 (*α*-H-4) and 5.33 (*β*-H-4)] and H-6 [δ_H 4.41, 4.27 (*α*-H-6), 4.43, 4.25 (*β*-H-6)] and HMBC correlation between H-3/C-7″, H-4/C-7‴ and H-6/C-7′. Thus, compound **33** was characterized as 3,4,6-tri-*O*-galloyl-D-glucose (Wilkins, 1988).

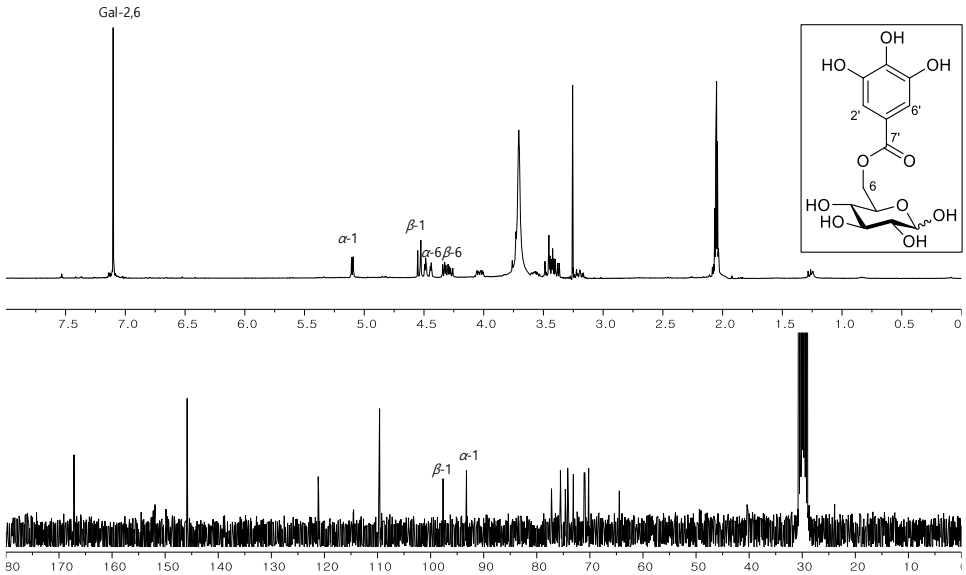


Figure 63. ^{1}H and ^{13}C NMR spectra of compound **31**

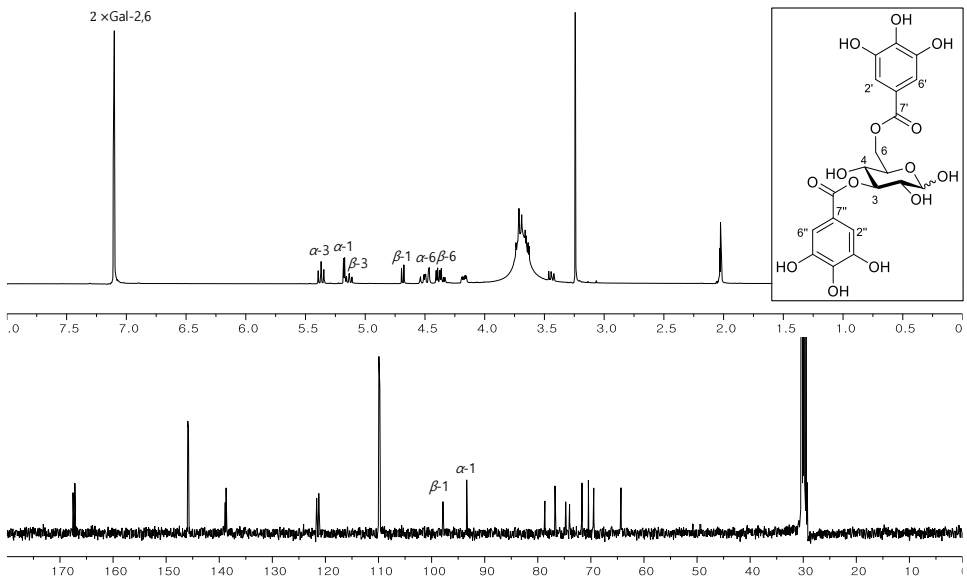


Figure 64. ^{1}H and ^{13}C NMR spectra of compound **32**

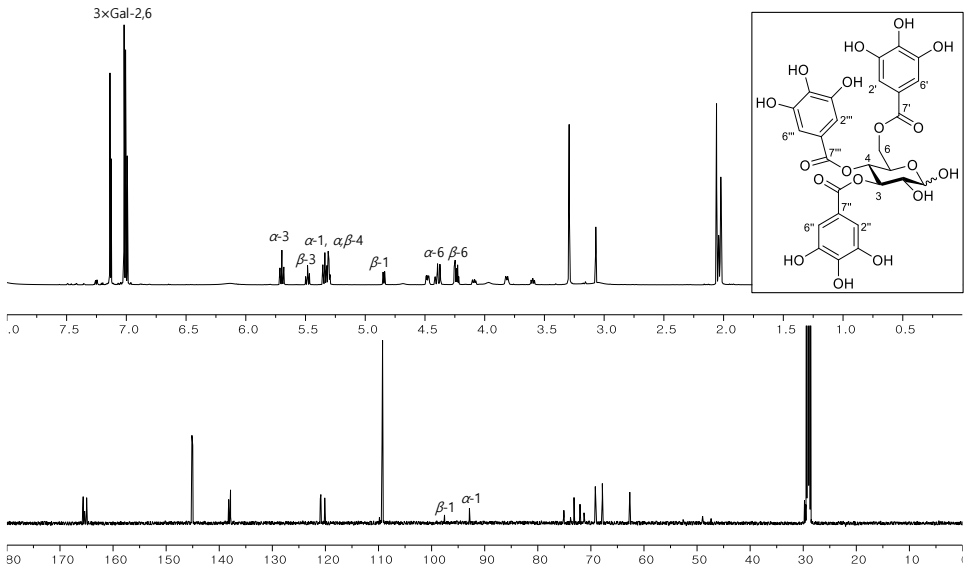


Figure 65. ^{1}H and ^{13}C NMR spectra of compound **33**

3.11. Compounds **34-37**

Compound **34** was obtained as colorless syrup and has the molecular formula $C_{14}H_{12}O_{11}$ according to negative HRESIMS analysis (*m/z*: 355.0302 $[M - H]^-$, calcd for $C_{14}H_{11}O_{11}$, 355.0301). The ^{13}C NMR data of **34** (Figure 66) displayed three

carboxylic signals [δ_C 174.4 (C-7), 172.9 (C-6), 170.8 (C-1)], one lactone signal (δ_C 163.3, C-7′), six aromatic signals [δ_C 145.3 (C-3′), 142.7 (C-5′), 138.7 (C-4′), 116.5 (C-6′**)**, 114.8 (C-1′), 107.6 (C-2′)], and four aliphatic signals [δ_C 76.3 (C-2), 43.7 (C-4), 35.6 (C-3), 34.1 (C-5)]. Furthermore, its ^{1}H NMR showed three methine protons at δ_H 5.08 (1H, *s*, H-2), 3.56 (1H, *d*, *J* = 8.4 Hz, H-3), 2.82 (1H, *m*, H-4) and one methylene at δ_H 2.67 (1H, *dd*, *J* = 11.3, 16.9 Hz, H-5a) and 2.06 (1H, *dd*, *J* = 3.5, 16.9 Hz, H-5b) with a discrete ^{1}H-^{1}H COSY system, and one-proton aromatic singlet at δ_H 6.87 (H-2′). The CD spectrum of **34** showed postivie Cotton effect at 235 nm. Based on above spectral data, compound **34** was identified as chebulic acid (Lee et al., 2007; Yoshida et al., 1982).

Compound **35** was obtained as yellowish syrup. (-)-HRESIMS analysis indicated that **35** had molecular formula of $C_{15}H_{14}O_{11}$ (*m*/*z* 369.0458 [M – H]$^-$, calcd for $C_{15}H_{13}O_{11}$, 369.0458). The ^{1}H and ^{13}C NMR spectra of **35** was analogous to those of **34** except for the signal of methoxy moiety. The HMBC analysis showed that the methoxy moiety was attached at C-6 (δ_C 173.9) by the correlation from δ_H 3.52 (3H, s, OMe). On the basis of above information, compound **35** was determined as 6-*O*-methyl chebulate (Li Xiang-yu et al., 2010).

Compound **36** was obtained as yellowish syrup and gave the molecular formula $C_{15}H_{14}O_{11}$ (*m*/*z* 369.0436 [M – H]$^-$, calcd for $C_{15}H_{13}O_{11}$, 369.0458). A comparison of the ^{1}H and ^{13}C NMR data and molecular formula of **36** suggested that compound **36** was regioisomer of **35**. The location of methoxy unit was determined to be C-1 by the HMBC correlation between δ_H 3.62 (3H, s, OMe) and δ_C 171.3 (C-1). Therefore, compound **36** was confirmed as 1-*O*-methyl chebulate (Li Xiang-yu et al., 2010).

Compound **37**, which had the same molecular formula as **34** according to its (-)-HRESIMS data (*m*/*z*: 355.0305 [M – H]$^-$, calcd for $C_{14}H_{11}O_{11}$, 355.0301), was

obtained as colorless syrup. Its 1D NMR spectra (Figure 69) was also analogous to that of **34**. Differences between **37** and **34** was the shielding of H-2 (δ_H 4.92, Δ-0.16 ppm) and deshielding of H-3 (δ_H 3.94, Δ0.28 ppm) and H-4 (δ_H 3.20, Δ0.38 ppm). Furthermore, the signal of C-5 (δ_C 31.5) was found to move to upfield. Based on above spectroscopic data, compound **37** was confirmed as neochebulic acid, which has the (2*S*, 3*S*, 4*R*) configurations, by comparison with literature (Ding et al., 2000).

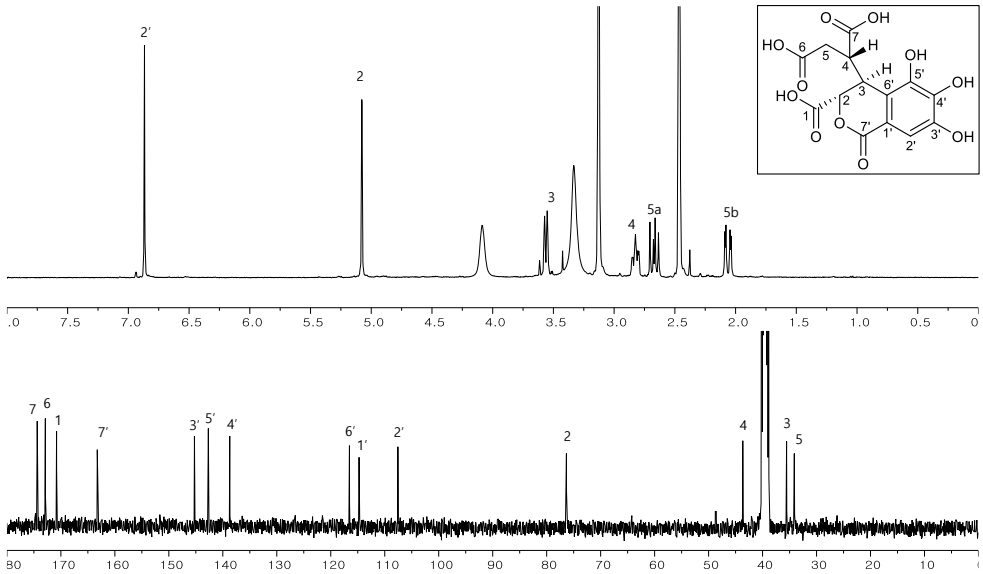


Figure 66. ^{1}H and ^{13}C NMR spectra of compound **34**

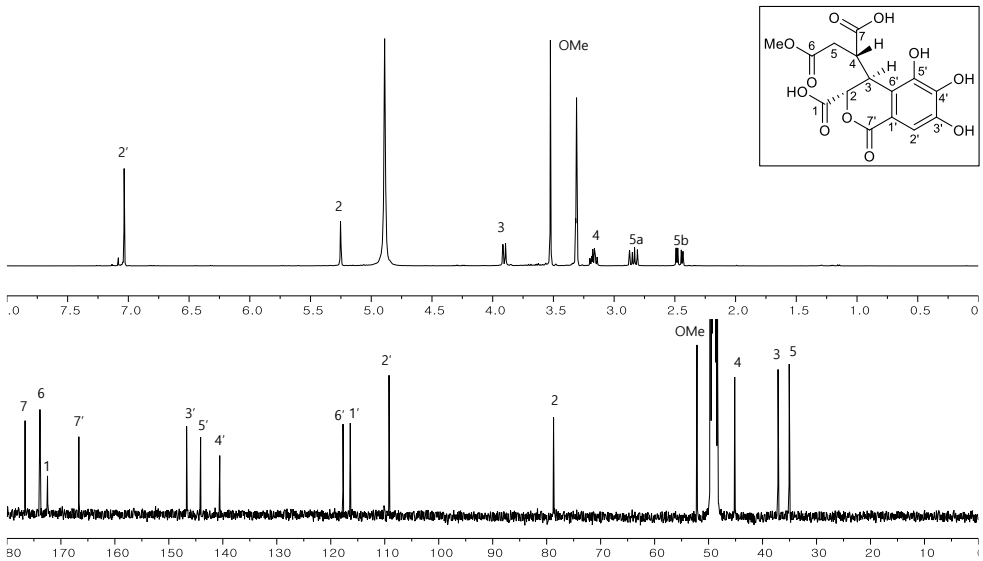


Figure 67. ^{1}H and ^{13}C NMR spectra of compound **35**

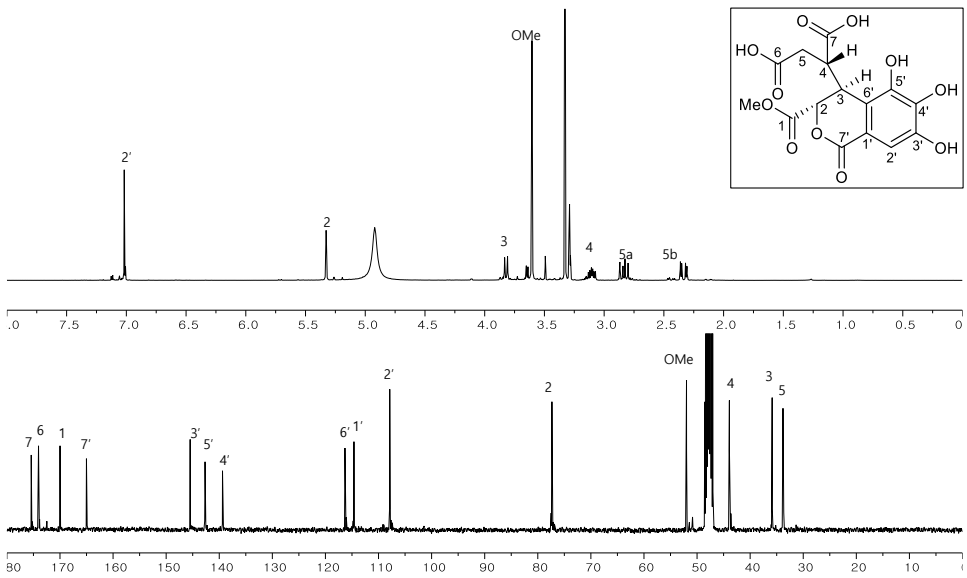


Figure 68. ^{1}H and ^{13}C NMR spectra of compound **36**

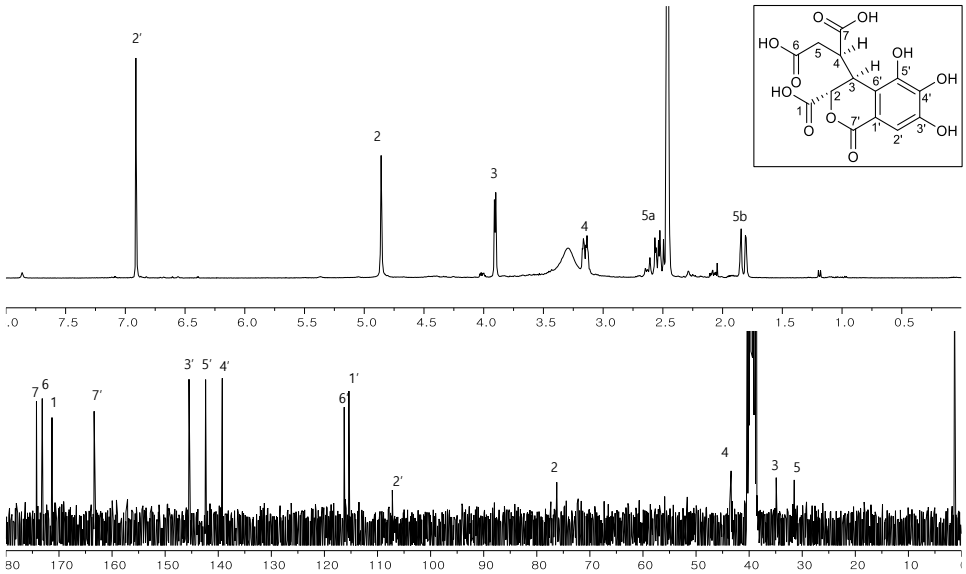


Figure 69. ^{1}H and ^{13}C NMR spectra of compound **37**

3.12. Compounds **38a** and **38b**

Compound **38a** and **38b** were obtained as white amorphous powders and equilibrium mixture in a ratio of 7:3. The molecular formula of **38a** and **38b** were both determined as $C_{14}H_{10}O_9$ by negative HRESIMS (321.0228 $[M - H]^-$, calcd for $C_{14}H_{10}O_9$, 321.0247). The MS fragment ion at *m*/*z* 169 [Gallic acid ($C_7H_6O_5$) – H]$^-$ and molecular formula of **38a** and **38b** indicated that their structures consisted of two galloyl groups. The ^{1}H NMR spectrum (**Figure 70**) showed two aromatic doublets with *meta*-coupling [δ_H 7.40 (1H, *d*, *J* = 1.8 Hz, *meta*-digallic acid-H-2), 7.26 (1H, *d*, *J* = 1.8 Hz, *meta*-digallic acid-H-6)] and three aromatic singlets at δ_H 7.23 (6/7 H, *s*, *para*-digallic acid-H-2', 6'), 7.22 (2H, s, *meta*-digallic acid-H-2', 6'), and 7.11 (6/7 H, *s*, *para*-digallic acid-H-2, 6). Based on above data and comparison with literature compound **38a** and **38b** was determined as *m*-digallic acid and *p*-digallic acid, respectively (Verzele et al., 1983).

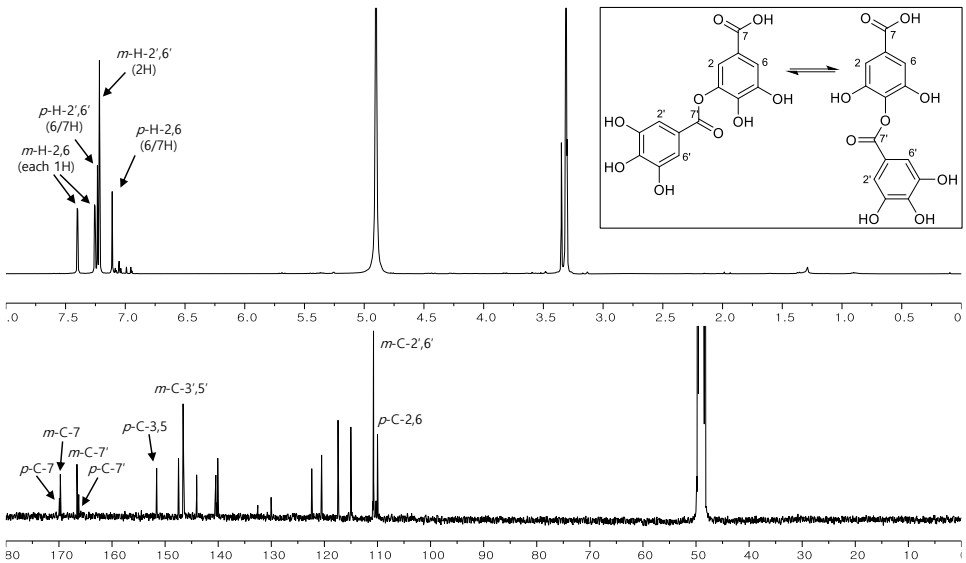


Figure 70. ^{1}H and ^{13}C NMR spectra of mixture of compound **38a** and **38b**

3.13. Compound **39**

Compound **39** was obtained as yellow amorphous powder and its molecular formula was determined as $C_{14}H_6O_8$ by the (-)-HRESIMS (*m*/*z*: 300.9983 $[M - H]^-$, calcd for $C_{14}H_6O_8$ 300.9984). The characteristic UV spectrum (λ_{max} 255 and 366 nm) suggested that compound **39** had a ellagic acid skeleton. The ^{1}H and ^{13}C NMR spectra (Figure 71) of **39** showed only one aromatic proton at δ_H 7.45 (2H, *s*, H-5, 5') and seven carbons at δ_C 159.1 (C-7, 7'), 148.1 (C-4, 4'), 139.6 (C-3, 3'), 136.3 (C-2, 2'), 112.3 (C-1, 1'), 110.2 (C-5, 5'), and 107.6 (C-6, 6'). Based on above information, compound **39** was considered to be symmetrical. On the basis of above information, compound **39** was confirmed as ellagic acid (Goriparti et al., 2013).

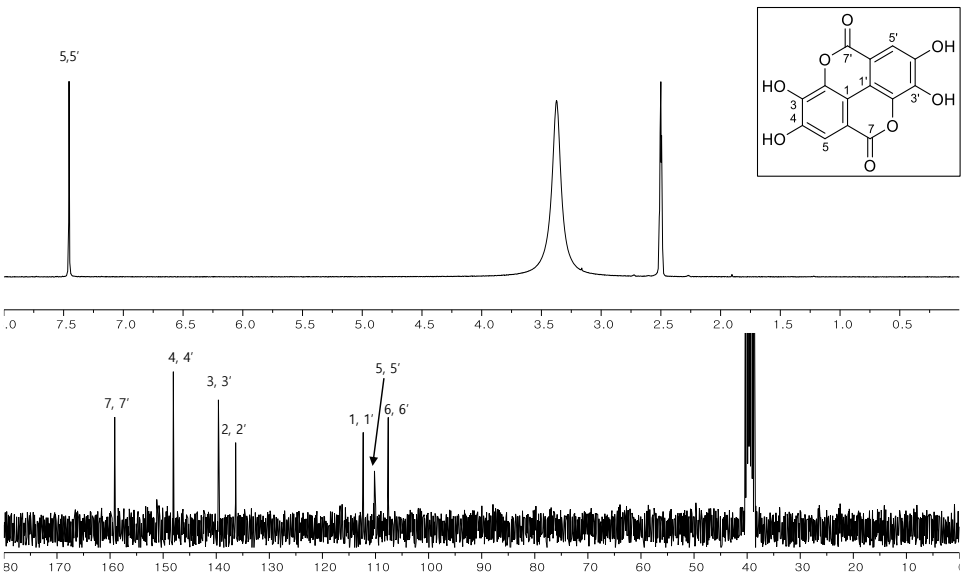


Figure 71. ^{1}H and ^{13}C NMR spectra of compound **39**

3.14. Compound **40**

Compound **40** was obtained as white amorphous powder and $[\alpha]_D^{20}$= 21.6 (c 0.10, MeOH). The molecular formula of **40** was $C_{13}H_8O_8$ as determined by combined (-)-HRESIMS (*m*/*z*: 291.0133 $[M - H]^-$, calcd for $C_{13}H_7O_8$, 291.0141) and ^{13}C NMR spectrum. Its ^{1}H NMR spectrum (Figure 72) showed one aromatic singlet (δ_H 7.27, 1H, *s*, H-3'), one methine [δ_H 4.36 (1H, *dd*, *J* = 2.0, 8.0 Hz, H-4)], and one methylene [δ_H 2.91 (1H, *dd*, *J* = 8.0, 18.5 Hz, H-5a) and 2.52 (1H, *m*, H-5b)]. In addition, its ^{13}C NMR spectrum displayed signals for one carbonyl [δ_C 193.6 (C-1)], one carboxylic [δ_C 173.4 (C-6)], one lactone [δ_C 160.3 (C-7')], two olefinic [δ_C 149.4 (C-2) and 145.2 (C-3)], six aromatic [δ_C 144.5 (C-6'), 140.6 (C-4'), 139.4 (C-5'), 115.2 (C-2'), 113.0 (C-1'), and 108.0 (C-3')] and two aliphatic carbons [δ_C 41.2 (C-4) and 37.5 (C-5)]. Based on above spectroscopic data, compound **40** was identified as brevifolincarboxylic acid by comparison with literature (Yoshida et al., 1992).

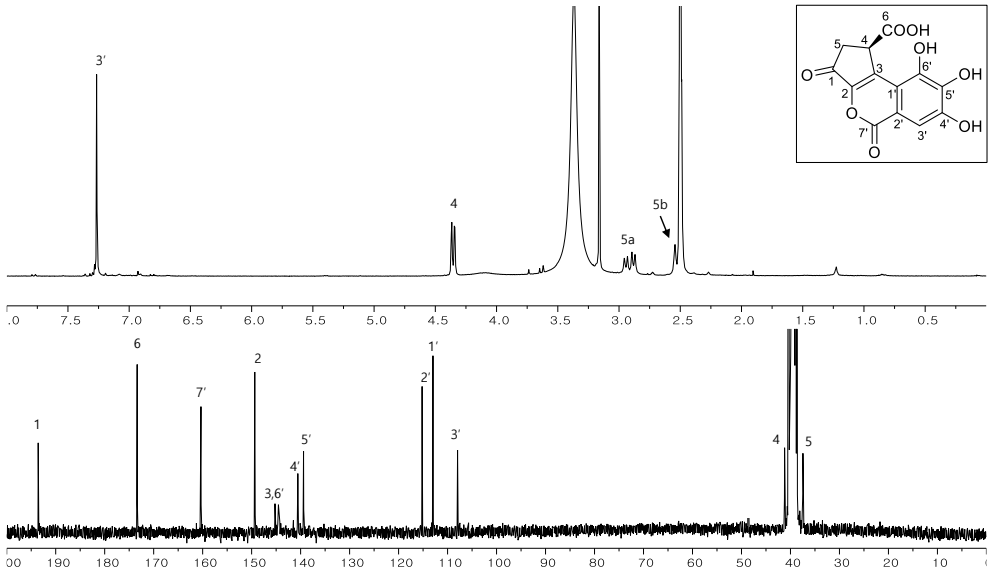


Figure 72. ^{1}H and ^{13}C NMR spectra of compound **40**

3.15. Compound **41**

Compound **41** was obtained as light brown amorphous powder and its molecular formula ($C_{27}H_{26}O_{20}$) was confirmed by the analysis of its (-)-HRESIMS data (*m/z*: 669.0923 $[M - H]^-$, calcd for $C_{27}H_{25}O_{20}$, 669.0916). The ^{1}H NMR spectrum (Figure 73) of **41** displayed signals at δ_H 7.18 (2H, *s*, Gal-2′,6′) indicating the presence of a galloyl unit. The large coupling constant of the H-1 signal (δ_H 5.69, 1H, d, J = 7.8 Hz) suggested the *β*-glucopyranose with $^{4}C_1$ conformation. Furthermore, the presence of neochebuloyl moiety was deduced from one-proton aromatic singlet at δ_H 7.18 (1H, *s*, Neocheb-H-2″), three methine protons at δ_H 5.31 (1H, *d*, *J* = 0.9 Hz, Neocheb-H-2′), 3.91 (1H, *d*, *J* = 8.4 Hz, Neocheb-H-3′) and 3.19 (1H, *ddd*, *J* = 4.2, 8.4, 11.1 Hz, Neocheb-H-4′), and one methylene at δ_H 2.90 (1H, *dd*, *J* = 11.1, 17.1 Hz, Neocheb-H-5′a) and 2.43 (1H, *dd*, *J* = 4.2, 17.1 Hz, Neocheb-H-5′b). The H-1 (δ_H 5.69) and H-6 (δ_H 4.33 and 3.92) signals of glucose were correlated to the galloyl carbonyl carbon at δ_C 165.2 (Gal-C-7′) and neochebuloyl carbonyl carbon at δ_C 166.3 (Neocheb-C-7′) in the HMBC spectrum, respectively. Thus, compound

41 was confirmed as phyllanemblinin F (Zhang et al., 2001).

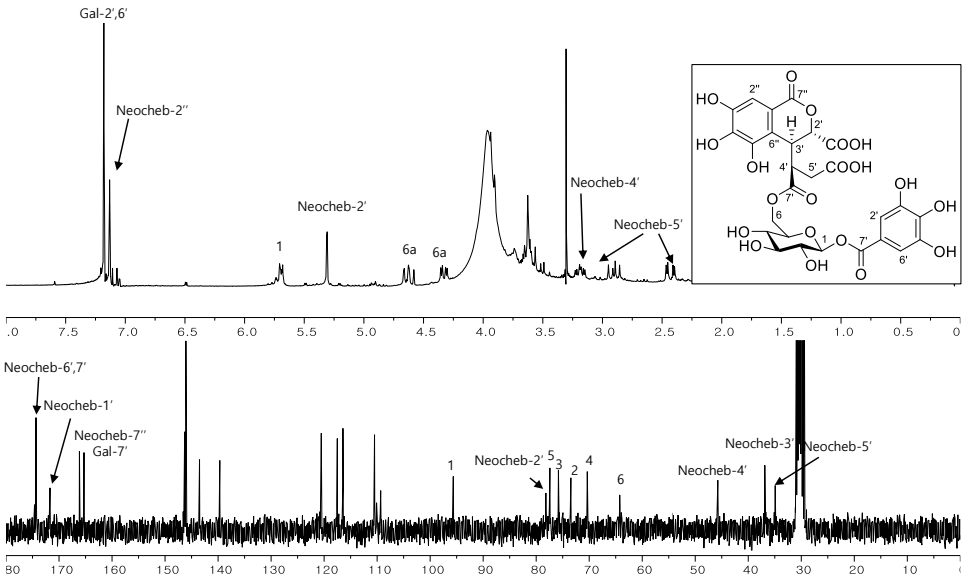


Figure 73. ^{1}H and ^{13}C NMR spectra of compound **41**

3.16. Compounds **42** and **43**

Compound **42** was obtained as yellowish amorphous powder. The molecular formula ($C_{27}H_{22}O_{18}$) was determined from (-)-HRESIMS (*m/z*: 663.0720 $[M - H]^-$, calcd for $C_{27}H_{21}O_{18}$, 663.0728) and ^{13}C NMR spectra. The $^{1}C_4$ conformation of the β-glucopyranose was deduced from the small coupling constants of the H-1 – H-4 [δ_H 6.36 (1H, *d*, *J* = 1.8 Hz, H-1), 4.80 (1H, br *s*, H-3), 4.46 (1H, *d*, *J* = 3.0 Hz, H-4), abd 3.98 (1H, *d*, *J* = 1.8 Hz, H-2)] signals and carbon signal of C-1 at δ_C 94.9 (Figure 74). Its ^{1}H NMR spectrum also displayed that the presence of one galloyl group [δ_H 7.05 (2H, *s*, Gal-H-2′, 6′)] and one HHDP group [δ_H 6.68 (1H, *s*, HHDP-H-5′), 6.66 (1H, *s*, HHDP-H-5″)]. The galloyl and HHDP group were determined to be attached to C-1 and C-3/C-6 of core glucose, respectively, by HMBC cross peak between H-1 (δ_H 6.36) and galloyl carbonyl carbon (δ_C 166.5) and H-3 (δ_H 4.80)/H-6 (δ_H 4.95 and 4.52) and HHDP carbonyl carbon signals (δ_C 168.4 and

170.0). Based on these spectral data, compound **42** was identified as corilagin (Luyen et al., 2014).

Compound **43** was obtained as light brown amorphous powder, and its molecular formula was identified as $C_{34}H_{26}O_{22}$ by the negative HRESIMS (*m*/*z*: 785.0840 [M – H]$^-$, calcd for $C_{34}H_{25}O_{22}$ 785.0837). Except for one additional galloyl moiety, its ^{1}H and ^{13}C NMR spectra were similar to those of compound **42 (**Figure 75**)**. The attachment of the one additional galloyl unit in the central glucose moiety was assigned to be the C-4 position by downfielded ^{1}H NMR signals of H-4 (δ_H 5.67) and HMBC correlation between the signals of H-4 and Gal-C-7″ (δ_C 166.3). Therefore, compound **43** was determined as tercatain (Tanaka et al., 1986b).

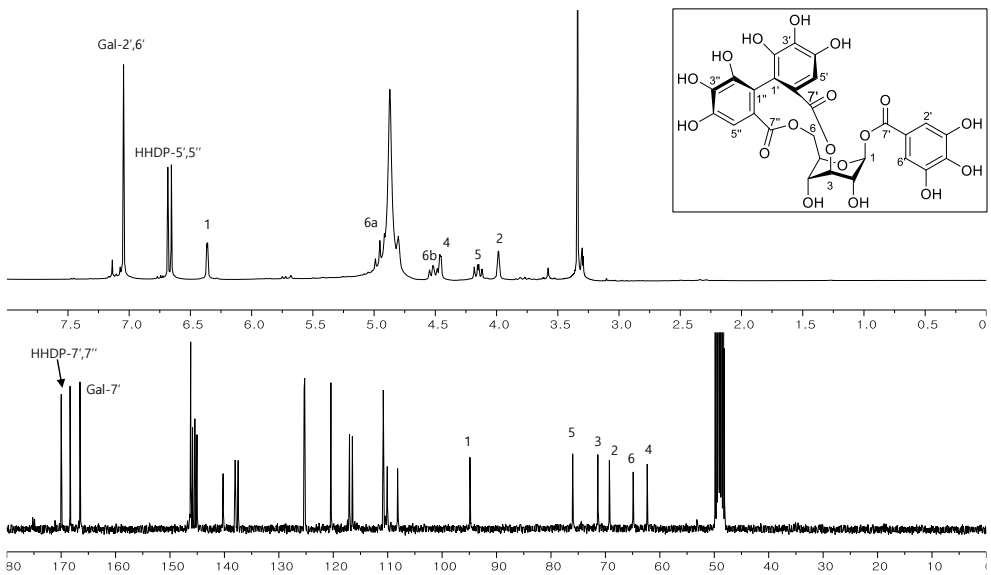


Figure 74. ^{1}H and ^{13}C NMR spectra of compound **42**

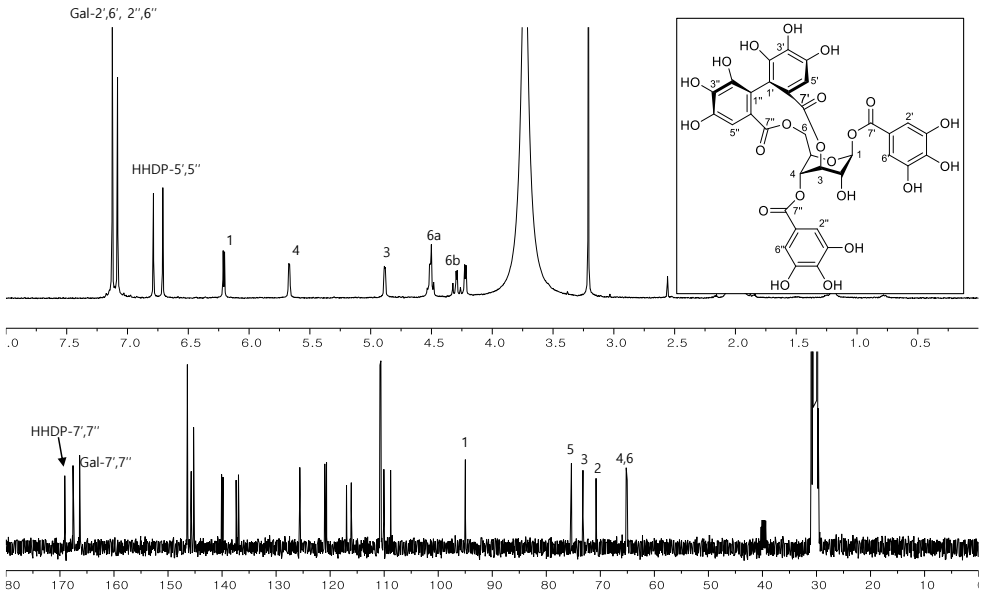


Figure 75. ^{1}H and ^{13}C NMR spectra of compound **43**

3.17. Compounds **44** and **45**

Compound **44** was obtained as light brown amorphous powder and equilibrium mixture in a ratio of 1:1. Its molecular formula was identified as $C_{27}H_{22}O_{18}$ by negative HRESIMS at *m*/*z* 633.0722 $[M - H]^-$ (calcd for $C_{27}H_{21}O_{18}$ 633.0728). The ^{1}H NMR signals at δ_H 5.25 (1H, *d*, *J* = 3.8 Hz, *α*-H-1) and 4.72 (1H, *d*, *J* = 7.8 Hz, *β*-H-1) and corresponding carbon signal at δ_C 93.7 (*α*-C-1) and 98.6 (*β*-H-1) indicated that compound **44** existed anomeric mixtures in solution (Figure 76). Furthermore, its ^{1}H NMR spectrum displayed one galloyl group [δ_H 7.02 (2H, *s*, *β*-Gal-H-2′,6′), 7.01 (2H, *s*, *β*-Gal-H-2′,6′)] and one HHDP group [δ_H 6.46, 6.45 (each 1H, *s*, *α*, *β*-HHDP-H-5′), 6.61, 6.60 (each 1H, *s*, *α*, *β*-HHDP-H-5′′)]. The location of a galloyl and HHDP moiety were determined to be C-3 and C-4/C-6 positions by HMBC correlation between H-3/Gal-C-7′, H-4/HHDP-C-7′ and H-6/HHDP-C-7′′, respectively. Thus, compound **44** was identified as gemin D by comparison with literature (Yoshida et al., 1985).

Compound **45** was obtained as light brown amorphous powder and equilibrium mixture in a ratio of 5:2. Its molecular formula $C_{34}H_{26}O_{22}$, indicated by (-)-HRESIMS (*m*/*z*: 785.0840 $[M - H]^-$, calcd for $C_{34}H_{25}O_{22}$ 785.0837). Its ^1H and ^{13}C NMR spectra indicated that compound **45** was also anomeric mitures (Figure 77). In addition, four aromatic singlets at δ_H 7.06, 6.98 (each 2H, *s*, *α*-Gal-H-2′,6′, 2″,6″), 7.05, 6.94 (each 0.8 H, *s*, *β*-Gal-H-2′,6′, 2″,6″) on its ^1H NMR spectrum evidenced the presence of one additional galloyl unit in compound **45** compared to **44**. The attachment of the additional galloyl unit was assigned to be the C-2 position by HMBC correlation between H-2/Gal-C-7″. Therefore, compound **45** was determined as tellimagrandin I (Hatano et al., 1988).

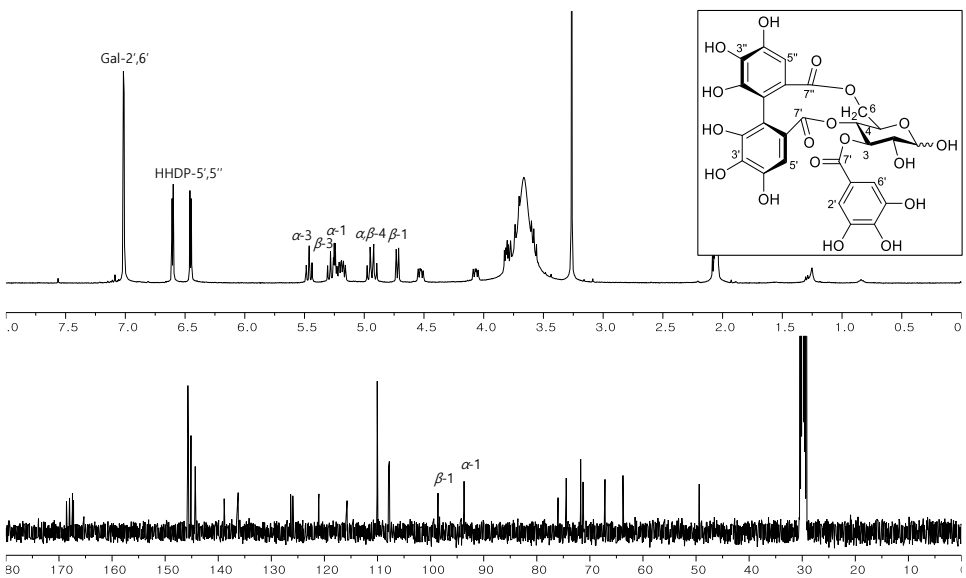


Figure 76. ^1H and ^{13}C NMR spectra of mixture of compound **44**

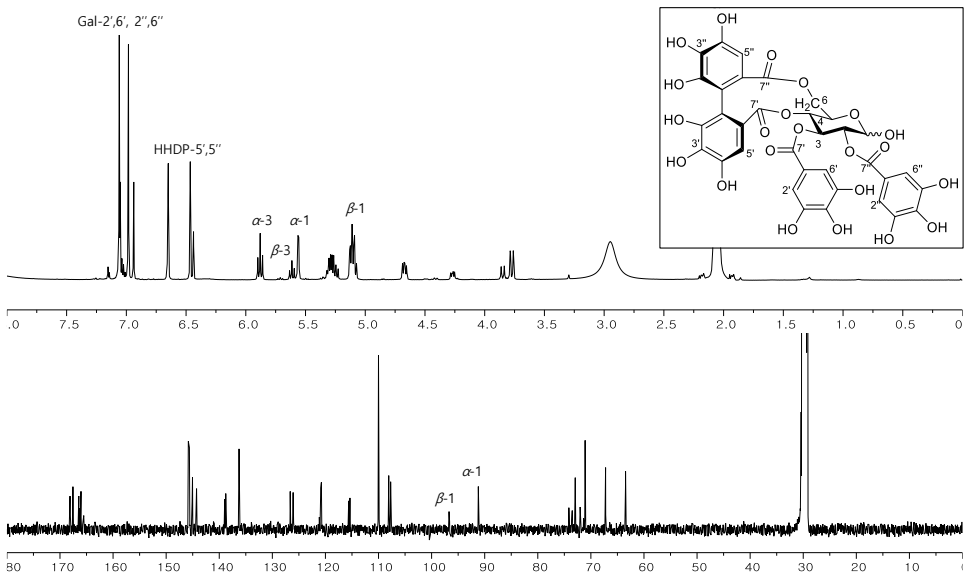


Figure 77. ^{1}H and ^{13}C NMR spectra of mixture of compound **45**

3.18. Compounds **46** and **47**

Compound **46** was obtained as yellowish amorphous powder. Its negative HRESIMS molecular ion at *m*/*z* 1083.0581 $[M - H]^-$ (calcd for $C_{43}H_{27}O_{30}$ 1083.0587) and ^{13}C NMR data were consistent with the molecular formula $C_{48}H_{28}O_{30}$. The MS fragment ion at *m*/*z* 601 [Gallagic acid ($C_{28}H_{10}O_{16}$) – H]$^-$ of **46** indicated that it contained gallagyl moiety (Mininel et al., 2014). Its ^{1}H spectrum (Figure 78) showed three aromatic one-proton singlets [δ_H 7.28 (1H, *s*, Gallagyl-H-2''), 6.70 (1H, *s*, Gallagyl-H-2'), 6.35 (1H, *s*, HHDP-H-5'')]. Since in ellagitannins, the aromatic proton signals are observed as a pair of singlet, the appearance of the odd number of aromatic singlets indicated the presence of a *C*-glycosidic linkage in its structure (Nonaka, 1988). The ^{13}C NMR spectrum of **46**, which showed six alipathic glucose carbons in the region δ_C 74.5 – 67.4 ppm, did not show the hemiacetal carbon which was expected at δ_C 90 – 100 ppm. These results and absence of HMBC correlation between H-1 (δ_H 5.41, 1H, *d*, *J* = 4.7 Hz) and C-5 (δ_C 72.8) indicate that the glucose moiety in **46** was in the open-chain form (Okuda

et al., 1983). The configuration of C-1 of open-chain glucose was determined to be *β*-configuration by coupling constant of H-1 (*d*, *J* = 4.7 Hz) and comparison with previous literature (Tanaka et al., 1986a). The HMBC analysis confirmed the positions of the C-glycosidic linkage *via* correlations from H-1 and H-2 (δ_H 4.87, *dd*, *J* = 2.2, 4.7 Hz) to HHDP-C-5' (δ_C 116.2) established the C-glycosidic linkage between C-1 and HHDP-C-5'. Furthermore, the location of HHDP and gallagyl moiety were determined to be C-2/C-3 [H-2 (δ_H 4.87)/HHDP-C-7' (δ_C 165.3), H-3 (δ_H 5.13)/HHDP-C-7" (δ_C 168.8)] and C-4/C-6 [H-4 (δ_H 4.23)/Gallagyl-C-7' (δ_C 168.0), H-6 (δ_H 3.81)/Gallagyl-C-7" (δ_C 167.9)], respectively, by analysis of HMBC spectrum. Based on above spectroscopic data, compound **46** was confirmed as punicacortein C (Tanaka et al., 1986a).

Compound **47** was obtained as light brown amorphous powder. The molecular formula of **47** was $C_{48}H_{28}O_{30}$ as determined by combined (-)-HRESIMS (*m*/*z*: 1083.0580 $[M - H]^-$, (calcd for $C_{43}H_{27}O_{30}$ 1083.0587) and ^{13}C NMR spectrum. The ^1H and ^{13}C NMR spectra (Figure 79) of **47** were analogous to those of **46**, but had some differences in the open-chain glucose moiety with the upfielded chemical shift and coupling constant of H-1 (δ_H 4.63, *s*) compared to those of **46**. According to previous literature, these characters indicated the C-1 of open-chain glucose was the *α*-configuration. Thus, compound **47** was identified as punicacortein D (Tanaka et al., 1986a).

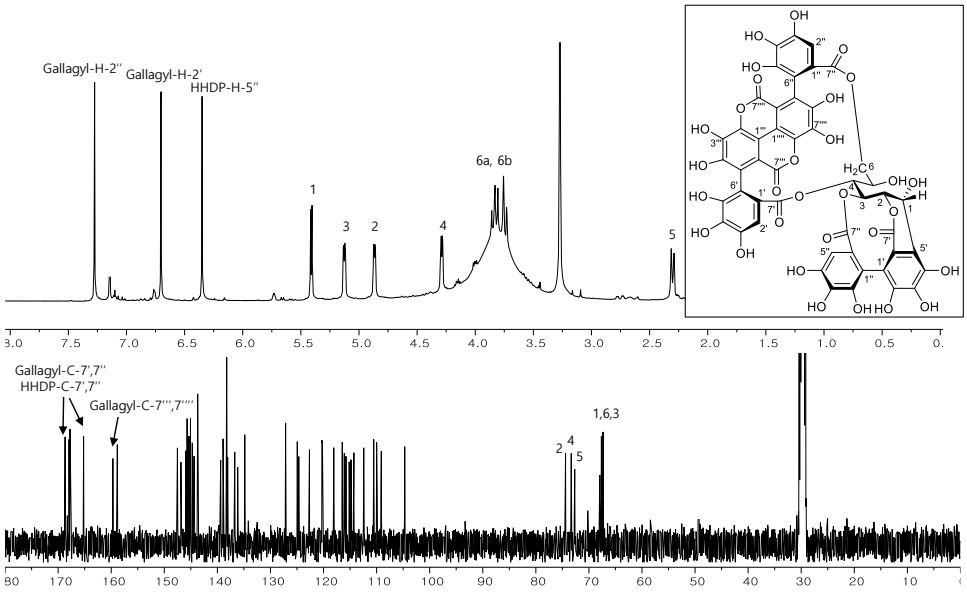


Figure 78. ^{1}H and ^{13}C NMR spectra of compound **46**

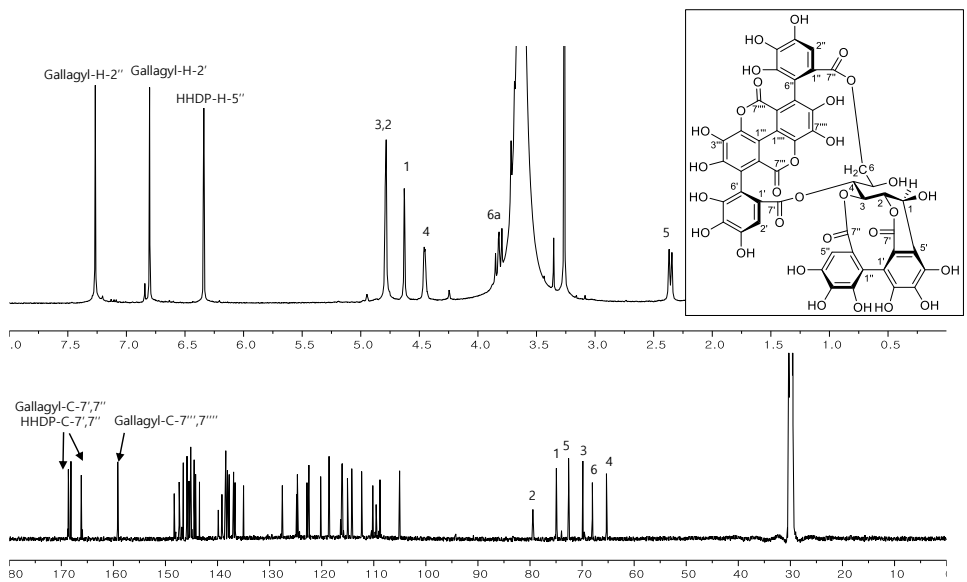


Figure 79. ^{1}H and ^{13}C NMR spectra of compound **47**

3.19. Compounds **48** and **49**

Compound **48** was obtained as light brown amorphous powder and equilibrium

mixture in a ratio of 5:3. Its molecular formula was identified as $C_{48}H_{28}O_{30}$ by negative HRESIMS at *m*/*z* 1083.0574 $[M - H]^-$ (calcd for $C_{48}H_{27}O_{30}$ 1083.0587). The ^1H NMR signals at δ_H 5.11 (1H, *d*, *J* = 3.5 Hz, *α*-H-1) and 4.72 (1H, *d*, *J* = 8.1 Hz, *β*-H-1) and corresponding carbon signal at δ_C 90.3 (*α*-C-1) and 94.4 (*β*-H-1) indicated that compound **48** existed anomeric mixtures in solution (Figure 80). The MS fragment ion at *m*/*z* 601 [Gallagic acid ($C_{28}H_{10}O_{16}$) – H]$^-$ and 301 [Ellagic acid ($C_{14}H_6O_8$) – H]$^-$ of **48** suggested that it contained gallagyl and HHDP moiety, respectively (Mininel et al., 2014). Its ^1H spectrum showed five aromatic singlets [δ_H 7.02 (1.6H, *s*, *α*, *β*-Gallagyl-H-2″), 6.65 (0.6H, s, *β*-HHDP-H-5″), 6.59 (0.6H, *s*, *β*-Gallagyl-H-2′), 6.58 (2H, s, *α*-HHDP-H-5", *α*-Gallagyl-H-2′), and 6.52 (1.6H, *s*, *α*, *β*-HHDP-H-5′)], which supported to the presence of gallagyl and HHDP moiety. The location of HHDP and gallagyl moiety were determined to be C-2/C-3 and C-4/C-6, respectively, by analysis of HMBC spectrum. Based on above spectroscopic data, compound **48** was confirmed as punicalagin (Ito et al., 2014).

Compound **49** was obtained as light brown amorphous powder and equilibrium mixture in a ratio of 5:3. Its molecular formula $C_{48}H_{30}O_{30}$, indicated by (-)-HRESIMS (*m*/*z*: 1085.0734 $[M - H]^-$, calcd for $C_{48}H_{29}O_{30}$ 1085.0744). Its ^1H and ^{13}C NMR spectra indicated that compound **49** was also anomeric mitures (Figure 81). The MS fragment ion at *m*/*z* 451 [flavogallonic acid ($C_{21}H_8O_{12}$) – H]$^-$ and 301 [Ellagic acid ($C_{14}H_6O_8$) – H]$^-$ of **49** suggested that it had flavogallonyl and HHDP moiety, respectively (Gu et al., 2013). Its ^1H NMR signals at δ_H 6.84 (2H, s, *α*-Gal-H-2′,6′) and 6.82 (1.2H, s, *β*-Gal-H-2′,6′) indicated the presence of a galloyl moiety. The attachment of the galloyl, flavogallonyl, and HHDP moieties were determined to be the C-6, C-4, and C-2/C-3 positions of central glucose by HMBC correlation, respectively. Therefore, compound **49** was identified as terflavin A by comparison with literature (Tanaka et al., 1986b).

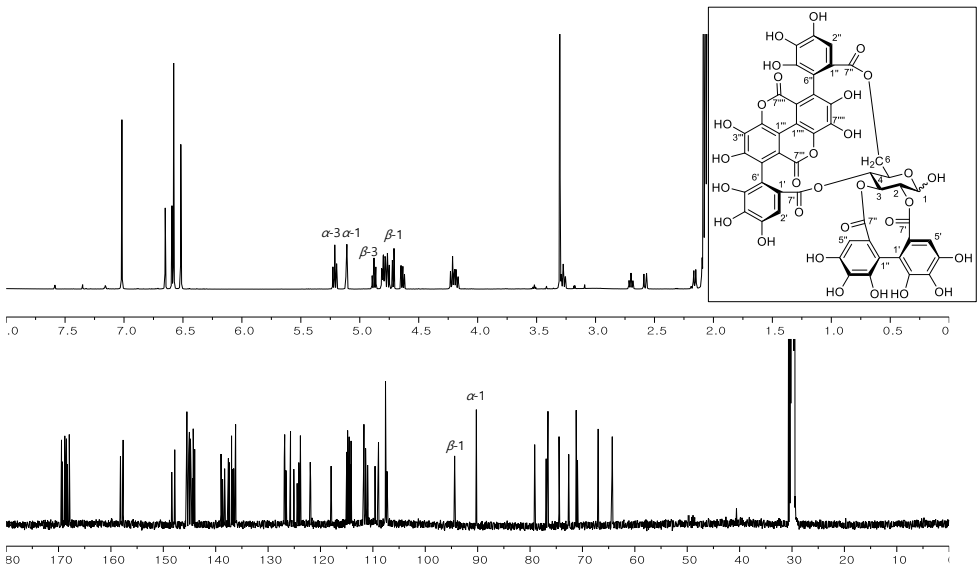


Figure 80. ^{1}H and ^{13}C NMR spectra of compound **48**

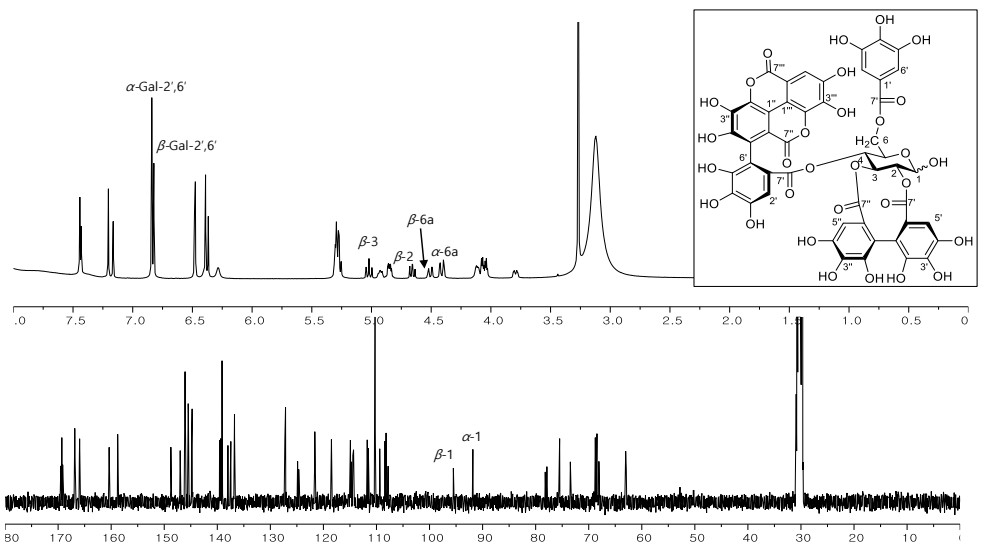


Figure 81. ^{1}H and ^{13}C NMR spectra of compound **49**

3.20. Compounds **50-55**

Compound **50** was obtained as white amorphous powder with the molecular formula $C_{30}H_{48}O_6$ as deduced from the negative HRESIMS, which showed a quasi-

molecular ion peak at *m*/*z* 503.3382 $[M - H]^-$ (calcd for $C_{30}H_{47}O_6$ 503.3373). Its ^1H NMR spectrum (**Figure 82**) displayed the presence of six tertiary methyls [δ_H 1.59 (H-27), 1.20 (H-29), 1.12 (H-30), 1.12 (H-24), 1.11 (H-26), and 1.09 (H-25)], one oxymethylene [δ_H 4.22 (1H, *d*, *J* = 10.5 Hz, H-23a) and 3.74 (1H, *d*, *J* = 10.5 Hz, H-23b)], three oxymethines [δ_H 4.28 (1H, *td*, *J* = 10.5, 4.1 Hz, H-2) 4.22 (1H, *d*, *J* = 9.1 Hz, H-3), and 3.61 (1H, br *s*, H-19)] and one olefinic protons [δ_H 5.57 (1H, br *t*, *J* = 4.1 Hz, H-12)]. The ^{13}C NMR spectrum exhibited signals for 30 carbons including one carboxyl (δ_C 181.4, C-28), one double bond (δ_C 145.4 and 123.5, C-13 and 12), three oxygenated methines [δ_C 81.6 (C-19), 78.7 (C-3), and 69.3 (C-2)], one oxygenated methylene (δ_C 66.9, C-23), and other 23 carbon signals had chemical shifts from 14.8 to 49.0. Based on these spectral data suggested that compound **50** had an olean-12-en-28-oic acid skeleton (Mahato and Kundu, 1994). The HMBC correlations indicated that four hydroxyl groups were existed at C-2, C-3, C-19, and C-23. Coupling constant between H-2 and H-3 (*J* = 9.1 Hz) indicated that the hydroxyl groups at C-2 and C-3 of **50** were both equatorially oriented (2*α*-, 3*β*-) (Bisoli et al., 2008). Based on these spectral data and comparison with literature, compound **50** was determined as arjungenin (Gossan et al., 2016).

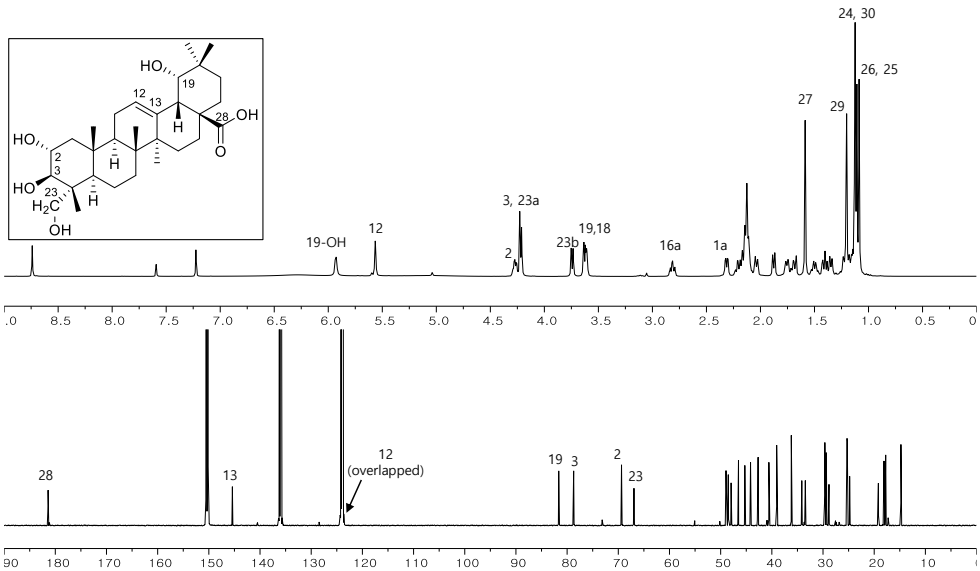


Figure 82. ^{1}H and ^{13}C NMR spectra of compound **50**

Compound **51** was obtained as a white amorphous powder and assigned a molecular formula of $C_{50}H_{68}O_{21}$ from the (-)-HRESIMS peak at *m/z* 1003.4206 $[M-H]^-$ (calcd for $C_{50}H_{67}O_{21}$ 1003.4175), which was confirmed by ^{13}C NMR spectrum. Extensive studies of the 1D and 2D NMR spectra of **51** indicated the presence of six tertiary methyl groups at δ_H 1.00 (H-30), 1.06 (H-24), 1.08 (H-25), 1.14 (H-29), 1.18 (H-26), and 1.89 (H-27) and one olefinic proton at δ_H 5.47 (br t, J = 4.2 Hz, H-12) with six corresponding methyl signals at δ_C 14.9 (C-24), 17.8 (C-25), 18.2 (C-26), 25.2 (C-30), 25.6 (C-27), 29.2 (C-29), two typical olefinic carbon signals at δ_C 123.9 (C-12) and 145.2 (C-13), indicating olean-12-ene as an aglycone (Figure 83). Furthermore, three oxygenated methine protons at δ_H 4.20 (1H, *m*, H-2), 4.11 (1H, *d*, J = 9.1 Hz, H-3), and 3.55 (1H, br *s*, H-19), and oxygen-bearing methylene protons at δ_H 4.92 and 4.41 (each 1H, *d*, J = 11.1 Hz, H-23) and the corresponding carbons [δ_C 69.1 (C-2), 78.0 (C-3), 81.6 (C-21), and 68.6 (C-23)] were given in the NMR spectra, suggesting the hydroxylation C-2, C-3, C-19, and C-23. This assumption was supported by the HMBC correlations (Figure 84) of δ_H 4.20 (H-2)

with δ_C 78.0 (C-3), δ_H 4.11 (H-3) with δ_C 14.9 (C-24), 43.6 (C-4), 68.6 (C-23), and 69.1 (C-2), δ_H 3.55 (H-19) with δ_C 25.5 (C-30), 29.2 (C-29), and 145.2 (C-13), and δ_H 4.92 and 4.41 (H-23) with 14.9 (C-24), 43.6 (C-4), 49.3 (C-5), and 78.0 (C-3). The orientations of the hydroxyl groups of C-2, C-3, and C-19 were assigned to be *α*, *β*, and *α*, respectively, according to the ROESY correlations [H-2 (δ_H 4.20) and H-25 (δ_H 1.08); H-3 (δ_H 4.11) and H-5 (δ_H 1.90); H-19 (δ_H 3.55) and H-30 (δ_H 1.00)] (Figure 85). The comparison of these 1D and 2D NMR data with the literature allowed the identification of the aglycone of **51** as the arjungenin (**50**, 2*α*,3*β*,19*α*,23-tetrahydroxyolean-12-en-28-oic acid) (Gossan et al., 2016). Furthermore, the presence of a *β*–D-glucopyranose moiety was deduced from an anomeric proton signal at δ_H 6.39 (d, J = 8.1 Hz, H-1‴) in the ^{1}H NMR spectrum and from the six signals at δ_C 96.3 (C-1‴), 74.7 (C-2‴), 79.5 (C-3‴), 71.6 (C-4‴), 79.8 (C-5‴), and 62.6 (C-6‴) in the ^{13}C NMR spectrum.(Agrawal, 1992) The location of this *β*–D-glucopyranose unit at C-28 was confirmed from the HMBC correlation between the proton at δ_H 6.39 (H-1‴) and C-28 at δ_C 177.8. In addition, the presence of one methylene at δ_H 3.75 (dd, J = 10.3 and 17.0 Hz, H-5′a) and 3.23 (dd, J = 3.7 and 17.0 Hz, H-5′b) and three methine protons at δ_H 6.00 (br s, H-2′), 5.03 (d, J = 8.3 Hz, H-3′) and 4.12 (m, H-4′) with a discrete ^{1}H-^{1}H COSY system and one-proton singlet at δ_H 7.89 (s, H-2″) on the ^{1}H NMR spectrum and carbon atoms of three carboxyl carbon signals at δ_C 175.0 (C-6′), 174.9 (C-7′), 173.5 (C-1′), one lactone signal at δ_C 165.6 (C-1″) on the ^{13}C NMR spectrum evidenced the presence of the neochebuloyl group which was converted into chebulic acid upon hydrolysis. This was also corroborated by the MS fragment ion at *m*/*z* 337 [(neochebuloyl ($C_{14}H_{12}O_{11}$) – H_2O – H]$^-$.(Pfundstein et al., 2010) The presence of the D-glucose and neochebuloyl moiety was confirmed by HPLC analysis on the acid hydrolysate of **51**. The absolute configuration of the neochebuloyl moiety was determined as (2′*S*, 3′*S*, 4′*S*) according to a positive Cotton effect at 235 nm in the

CD spectrum of chebulic acid, which was isolated from acid hydrolysate of **1**.(Yoshida et al., 1982) Thus, compound **51** was characterized as 23-*O*-neochebuloylarjungenin 28-*O*-*β*–D-glucopyranosyl ester.

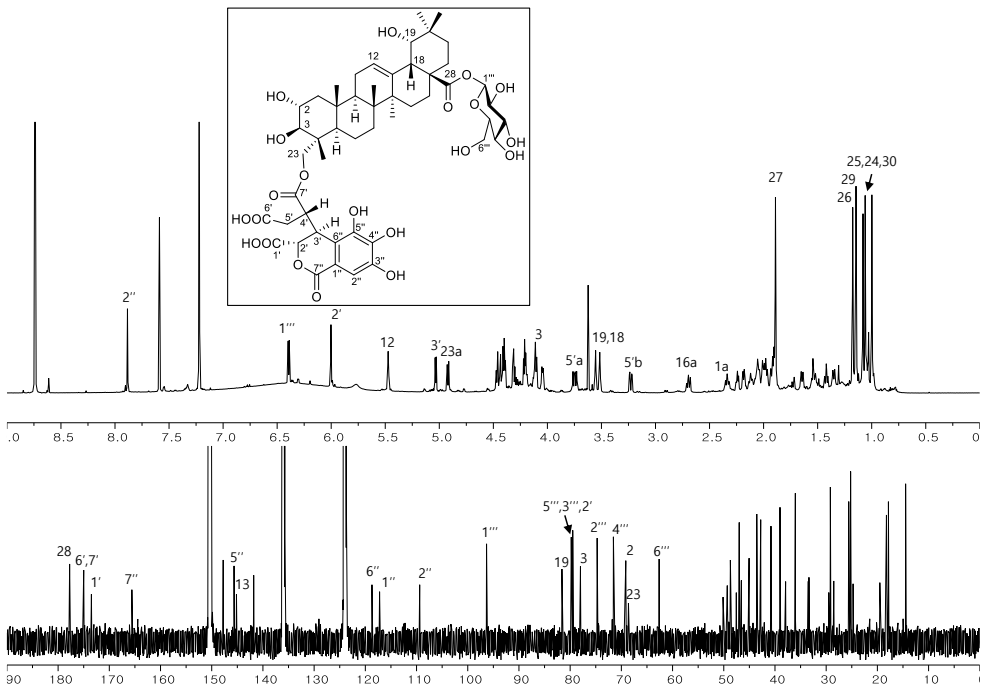


Figure 83. ^{1}H and ^{13}C NMR spectra of compound **51**

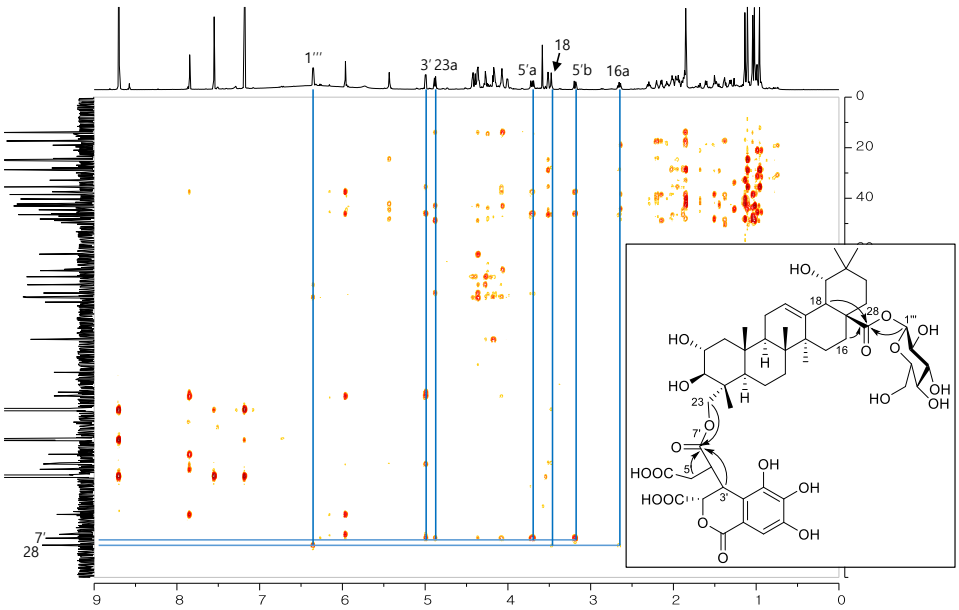


Figure 84. HMBC spectrum of compound **51**

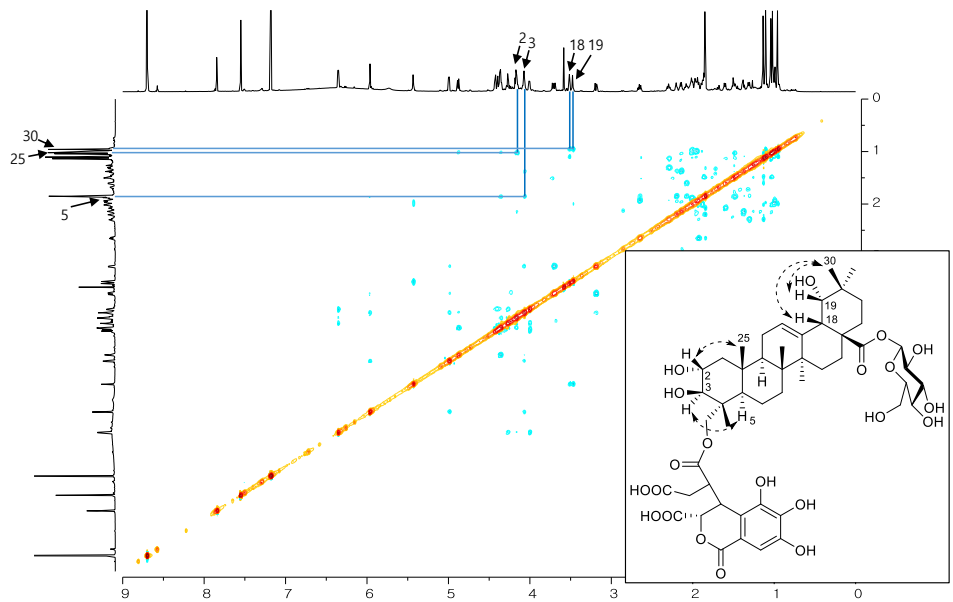


Figure 85. ROESY spectrum of compound **51**

Compound **52** was isolated as a light brown, amorphous powder. Its (-)-HRESIMS peak at *m/z* 841.3647 $[M-H]^-$ (calcd for 841.3647) established a

molecular formula of $C_{44}H_{58}O_{16}$, which is 162 mass units ($C_6H_{10}O_5$) less than that of **51**, suggesting that **52** is a deglycosylated derivative of **51**. Except for the signals corresponding to the glucose moiety, the ^{1}H and ^{13}C NMR spectra (Figure 86) of **52** were very similar to those of **51**, but had some differences in the neochebuloyl moiety with the shielding of H-2′ (δ_H 5.80, Δ-0.20 ppm) and deshielding of H-3′ (δ_H 5.32, Δ+0.29 ppm) and H-4′ (δ_H 4.52, Δ+0.40 ppm) compared to those of **51** (Figure 86). In addition, the coupling constant of H-3′ (*d*, *J* = 4.0 Hz) was smaller than that of **51** (*d*, *J* = 8.3 Hz). Furthermore, the absence of ROESY cross peaks between H-2′ and H-4′ provided further evidence for the stereochemistry of C-4′ (Figure 89). These spectroscopic differences between **52** and **51** were analogous to those between the chebulic acid (**34**) and neochebulic acid (**37**) with the 2′*S*, 3′*S*, and 4′*R* configurations. Furthermore, acid hydrolysate of **52** yielded neochebulic acid identified by HPLC analysis and comparison of CD spectra (positive Cotton effect at 213 nm) with an authentic sample. Thus, compound **2** was elucidated as 23-*O*-4′-*epi*-neochebuloylarjungenin.

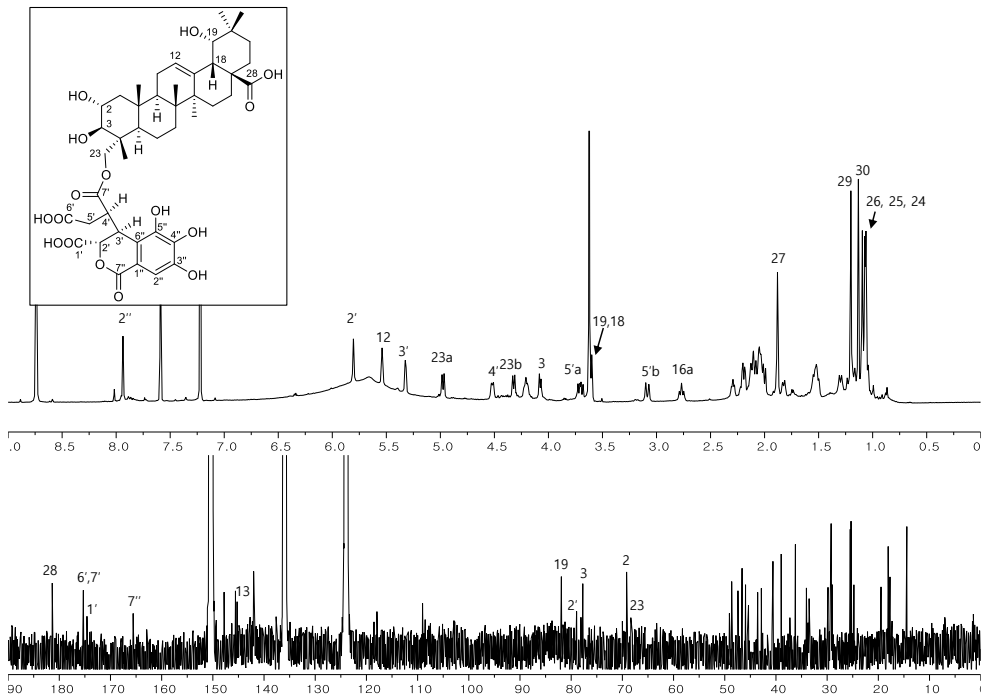


Figure 86. ^{1}H and ^{13}C NMR spectra of compound **52**

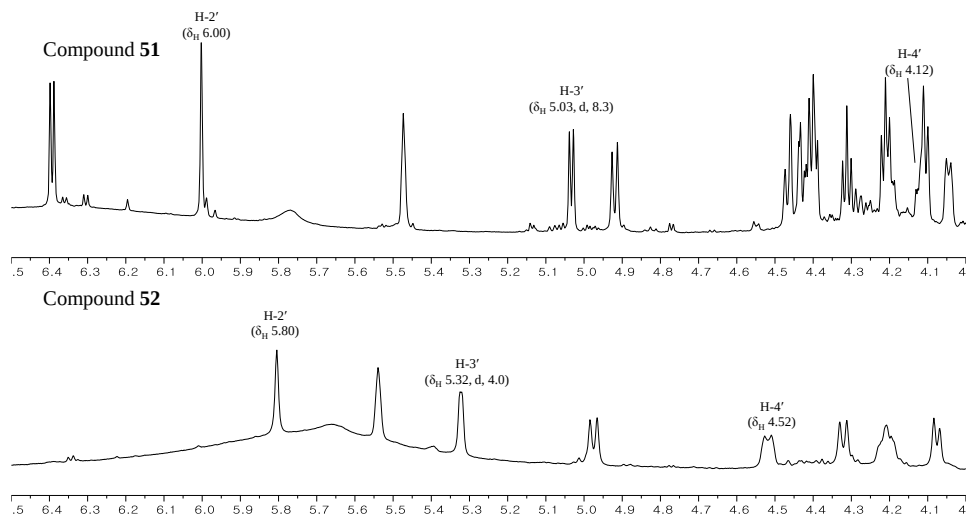


Figure 87. Comparison of expanded ^{1}H NMR spectra between **51** and **52**

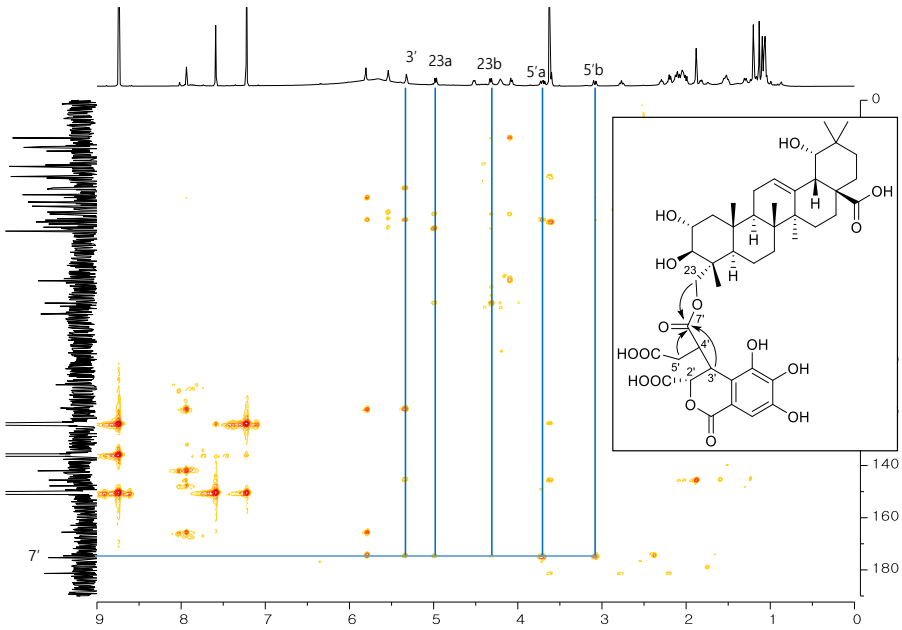


Figure 88. HMBC spectrum of compound **52**

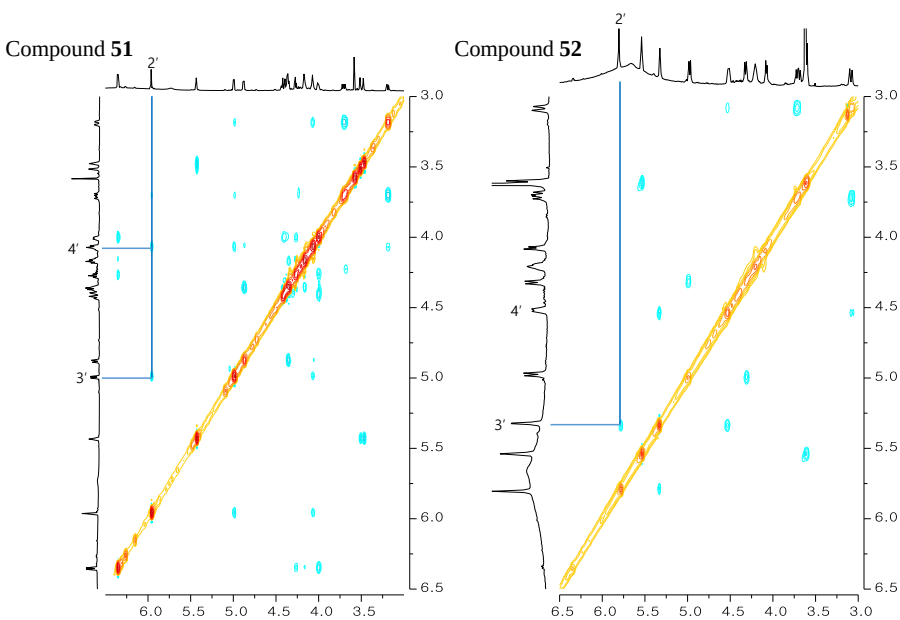


Figure 89. Comparison of expanded ROSEY spectra between **51** and **52**

Compound **53** was obtained as white amorphous powder and its molecular

formula was determined as $C_{37}H_{52}O_{10}$ by (-)-HRESIMS (*m/z*: 655.3484 $[M - H]^-$, calcd for $C_{37}H_{51}O_{10}$, 655.3482). Its ^1H and ^{13}C NMR were analogous to those of compound **50**, except for signals at δ_H 7.89 (2H, s, H-2′,6′) and δ_C 167.7 (C-7′), 148.2 (C-3′,5′), 141.5 (C-4′), 122.0 (C-1′), and 110.6 (C-2′,6′) corresponding to galloyl group (Figure 90). The location of galloyl group was determined to be the C-23 positions by strongly downfielded ^1H NMR signals of H-23 (δ_H 4.74 and 4.51, each 1H, *d*, *J* = 11.1 Hz) and HMBC correlation between the signals of H-23 and C-7′. Thus compound **53** was identified as 23-*O*-galloylarjunic acid (Machumi et al., 2013).

Compound **54** was obtained as brown amorphous powder. Quasi-molecular ion peak of (-)-HRESIMS at *m/z* 817.4030 $[M - H]^-$ (calcd for $C_{43}H_{61}O_{15}$ 817.4010) indicated the molecular formula of **54** was $C_{43}H_{62}O_{15}$. The ^1H and ^{13}C NMR spectra (Figure 91) of **54** were similar to those of **53**, but there was an additional *β*-glucose moiety in structure of **54**, which was deduced from a anomeric proton at δ_H 6.38 (1H, *d*, *J* = 8.2 Hz, H-1″) and anomeric carbon signal at δ_C 96.4 (C-1"). The location of *β*–D-glucopyranose unit at C-28 was confirmed from upfielded carbon signal of C-28 (δ_C 177.8) and HMBC correlation between H-1″ and C-28. Thus, compound **54** was confirmed as quercotriterpenoside I (Marchal et al., 2011).

Compound **55** was obtained as white amorphous powder and (-)-HRESIMS determined its molecular formula as $C_{36}H_{58}O_{11}$, which was indicated by *m/z* value of 711.3958 $[M + HCOOH - H]^-$ (calcd for $C_{37}H_{59}O_{13}$ 711.3956). Based on molecular formula and NMR spectra, compound **55** was considered to be a degalloylated derivative of **54** (Figure 92). The attachment of the *β*–D-glucopyranose moiety was assigned to be the C-28 position by upfielded carbon signal of C-28 (δ_C 177.8) and HMBC correlation between H-1′ (δ_H 6.39) and C-28. Thus, compound **55** was identified as arjunglucoside I (Abe and Yamauchi, 1987)

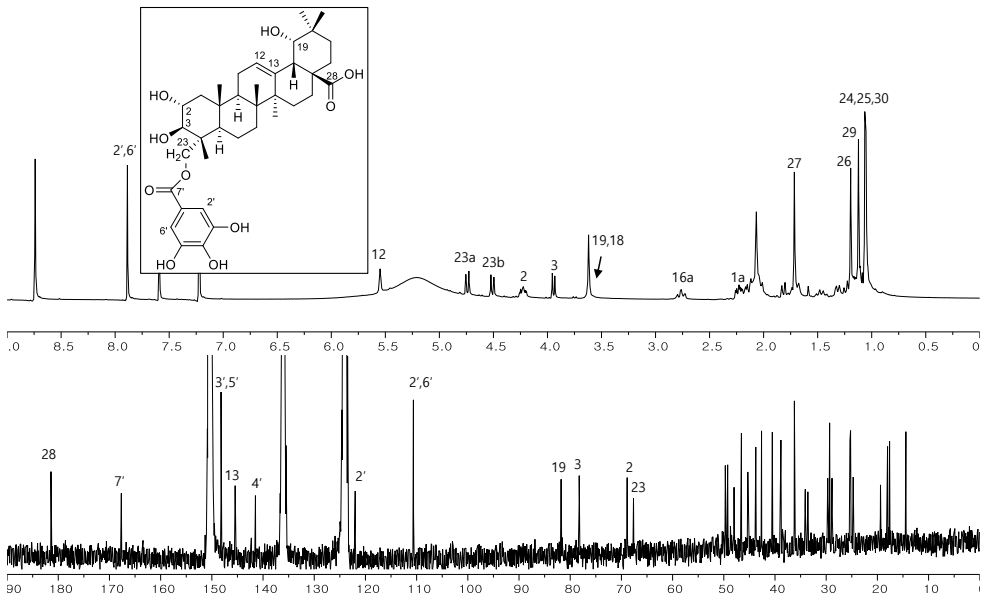


Figure 90. ^{1}H and ^{13}C NMR spectra of compound **53**

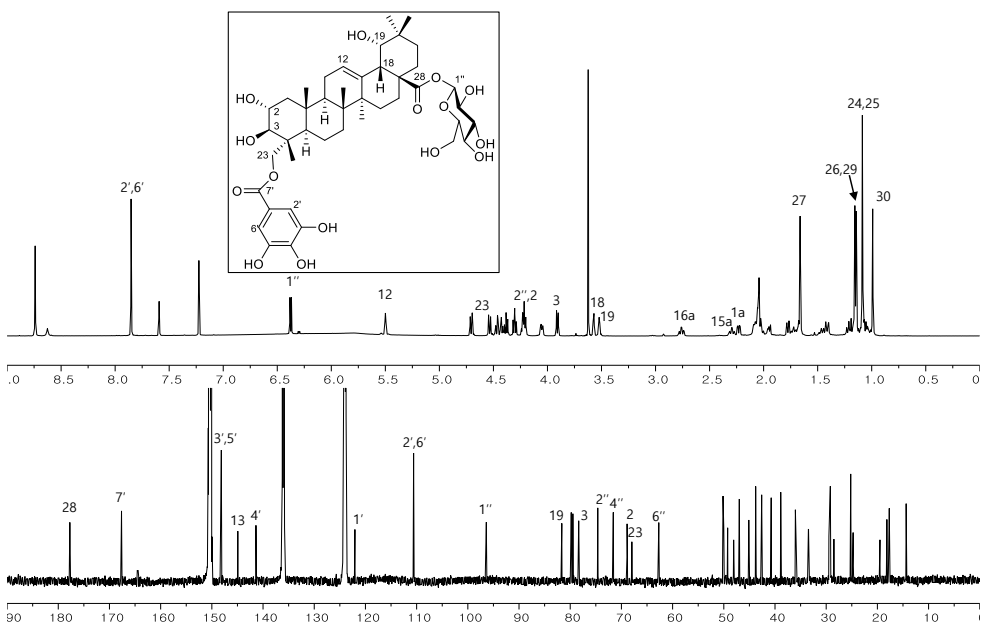


Figure 91. ^{1}H and ^{13}C NMR spectra of compound **54**

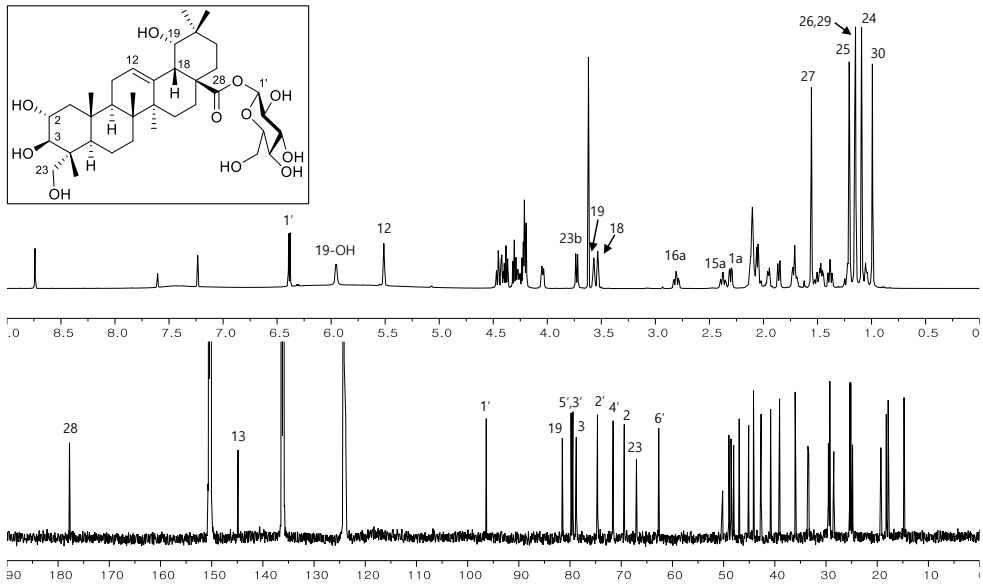


Figure 92. ^{1}H and ^{13}C NMR spectra of compound **55**

3.21. Compounds **56** and **57**

Compound **56** was obtained as white amorphous powder. The molecular formula was determined to be $C_{30}H_{48}O_6$ from (-)-HRESIMS at *m*/*z* 503.3384 $[M - H]^-$ (calcd for $C_{30}H_{47}O_6$ 503.3373). The ^{1}H NMR spectrum (Figure 93) of **56** exhibited six tertiary methyl groups [δ_H 1.80 (3H, *s*, H-25), 1.76 (3H, *s*, H-24), 1.64 (3H, *s*, H-26), 1.25 (3H, *s*, H-27), 1.02 (3H, *s*, H-30), 0.94 (3H, *s*, H-29)] one olefinic proton [δ_H 5.58 (1H, br *s*, H-12)], one oxymethylene [δ_H 4.42 (1H, *d*, *J* = 10.4 Hz, H-23a), 4.06 (1H, *d*, *J* = 10.4 Hz, H-23b)], and three oxymethines [δ_H 5.11 (1H, br *s*, H-6), 4.43 (1H, *m*, H-2), 4.25 (1H, *d*, *J* = 9.3 Hz, H-3)]. In additions, its ^{13}C NMR spectrum showed six tertiary methyl groups [δ_C 33.7 (C-29), 26.8 (C-27), 24.3 (C-30), 19.5 (C-25), 19.1 (C-26), 16.4 (C-24)] four oxygenated carbons [δ_C 78.9 (C-3), 69.6 (C-2), 68.1 (C-6), 66.7 (C-23)], two olefinic carbons [δ_C 144.7 (C-13), 123.4 (C-12)], and one carboxylic carbon [δ_C 180.7 (C-28)]. These spectral data indicated that compound 56 was a tetra-hydroxyolean-12-enoic acid (Mahato and Kundu, 1994). The comparison of these spectral data with the previous study allowed the

compound **56** as terminolic acid (2*α*,3*β*,6*β*,23-tetrahydroxyolean-12-en-28-oic acid) (Li et al., 2002).

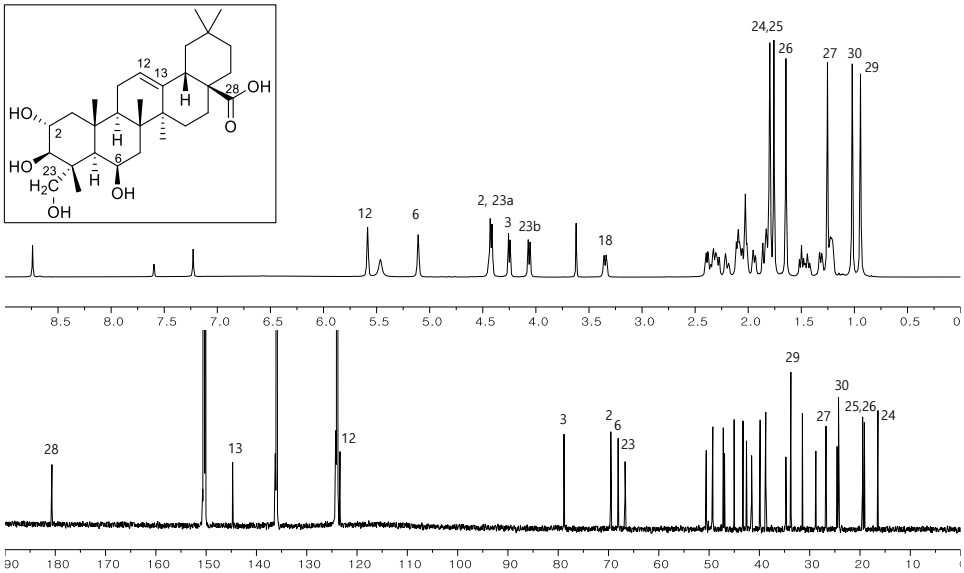


Figure 93. ^{1}H and ^{13}C NMR spectra of compound **56**

Compound **57** was obtained as light brown amorphous powder. Its (-)-HRESIMS showed a quasi-molecular ion peak at *m/z* 817.4025 $[M - H]^-$ (calcd for $C_{43}H_{61}O_{15}$ 817.4010) consistent with the formula of $C_{43}H_{62}O_{15}$. Analysis of its ^{1}H and ^{13}C NMR spectra (Figure 94) showed the presence of one *β*-glucopyranosyl moiety from signal at δ_H 6.29 (1H, *d*, *J* = 7.9 Hz, H-1") and an anomeric carbon at δ_C 96.3 (C-1"). Furthermore, the presence of galloyl unit was deduced from one aromatic singlet of two protons at δ_H 7.84 (H-2',6'), and one carboxyl ester carbon at δc 167.7 (C-7′), three oxygenated aromatic carbons at δc 148.1 (C-3′ and 5′) and 141.5 (C-4′), one aromatic quaternary carbon at δc 121.9 (C-1′), and two aromatic methines at δ_C 108.5 (C-2′ and 6′). Except for the signals corresponding to the *β*-glucopyranosyl and galloyl moiety, the ^{1}H and ^{13}C NMR spectra of **57** was analogous to those of **56**. The locations of the *β*-glucopyranosyl and galloyl moiety

was determined to be C-28 (δc 176.9) and C-23 (δc 67.6) position of aglycone by analysis of its HMBC spectrum. With above observed spectroscopic data, compound **57** was identified as 23-*O*-galloylterminolic acid 28-*O*-β-D-glucopyranosyl ester.

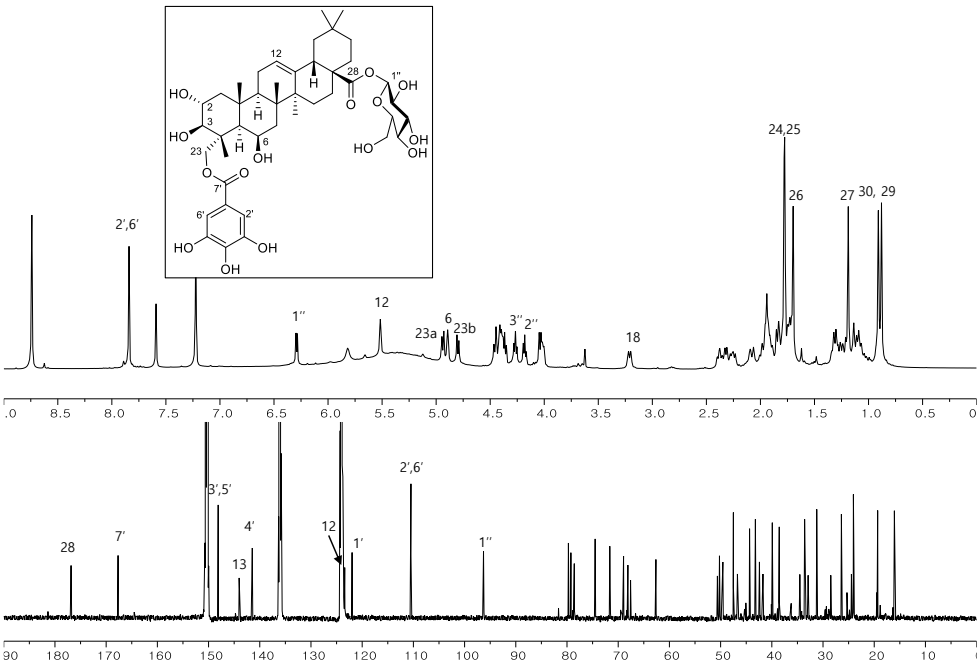


Figure 94. ^{1}H and ^{13}C NMR spectra of compound **57**

3.22. Compounds **58**-**61**

Compound **58** was obtained as white amorphous powder with molecular formula $C_{30}H_{48}O_5$ indicated by (-)-HRESIMS (*m*/*z*: 487.3436 $[M - H]^-$, calcd for $C_{30}H_{47}O_5$ 487.3423). The ^{1}H NMR spectrum (Figure 95) of **58** exhibited the presence of six tertiary methyls groups [δ_H 1.22 (H-27), 1.11 (H-30), 1.09 (H-24), 1.09 (H-26), 1.06 (H-25), and 0.93 (H-29)], one olefinic protons [δ_H 5.48 (1H, br *t*, *J* = 3.2 Hz, H-12)], one oxymethylene [δ_H 4.24 (1H, *d*, *J* = 10.5 Hz, H-23a) and 3.74 (1H, *d*, *J* = 10.5 Hz, H-23b)], two oxymethines [δ_H 4.25 (1H, *m*, H-2) 4.22 (1H, *d*, *J* = 8.2 Hz, H-3)]. The ^{13}C NMR spectrum showed signals for 30 carbons including one

carboxyl (δ_C 180.7, C-28), one double bond (δ_C 145.3 and 123.0, C-13 and 12), two oxygenated methines [δ_C 78.6 (C-3), and 69.4 (C-2)], one oxygenated methylene (δ_C 66.9, C-23), and other 24 carbon signals had chemical shifts from 14.9 to 48.7. Based on these spectral data indicated that compound **58** was trihydroxy-olean-12-en-18-oic acid skeleton. The locations of three hydroxyl groups were determined to be at C-2, C-3, and C-23 from its HMBC spectrum. Furthermore, large coupling constant between H-2 and H-3 (J = 8.2 Hz) suggested to their axial disposition, which indicated that the hydroxyl groups at C-2 and C-3 of **58** were both equatorially (2α-, 3β-) oriented. Based on these results, compound **58** was determined as arjunolic acid by comparison with literature (Huang et al., 2016).

Compound **59** was obtained as white amorphous powder and displayed a formic acid adducted ion at m/z 695.4022 $[M + HCOOH - H]^-$ (calcd for $C_{37}H_{59}O_{12}$ 695.4007) in the (-)-HRESIMS consistent with the molecular formula $C_{36}H_{58}O_{10}$. The ^1H and ^{13}C NMR of **59** were similar to those of compound **58**, except for signals at δ_H 6.34 (1H, *d*, J = 8.1 Hz, H-1′) and six oxygenate protons between δ_H 4.49 and 4.05 ppm and δ_C 96.3 (C-1′), 79.9 (C-5′), 79.4 (C-3′), 74.6 (C-2′), 71.7 (C-4′) and 62.8 (C-6′) corresponding to β-glucopyranosyl moiety (Figure 96). The location glucopyranose moiety of was determined to be the C-28 position by upfielded ^{13}C NMR signal of C-28 (δ_C 177.0) and HMBC correlation between the signals of H-1′ and C-28. Thus compound **53** was confirmed as arjunglucoside II (Bisoli et al., 2008).

Compound **60** was obtained as white amorphous powder. Its molecular formula of $C_{37}H_{52}O_9$ was established by the quasi-molecular ion peak at m/z 639.3535 $[M - H]^-$ (calcd for $C_{37}H_{51}O_9$ 639.3535). The ^1H and ^{13}C NMR of **60** were analogous to those of compound **58**, but additional signals at δ_H 7.85 (2H, *s*, H-2′,6′) and δ_C 167.6 (C-7′), 148.2 (C-3′,5′), 141.5 (C-4′), 122.0 (C-1′), and 110.6 (C-2′,6′),

corresponding to galloyl moiety, were shown (Figure 97). The location of galloyl group at C-23 was suggested by the pronounced downfield shift of H-23 (δ_H 4.67 and 4.50, each 1H, *d*, *J* = 11.0 Hz) and HMBC correlation between the signals of H-23 and C-7′. Thus compound **53** was identified as 23-*O*-galloylarjunolic acid (Conrad et al., 1998).

Compound **61** was obtained as light brown amorphous powder. Its HRESIMS (negative mode) showed an ion peak at *m*/*z* 801.4066 $[M - H]^-$ (calcd for $C_{43}H_{61}O_{14}$ 801.4061) consistent with the molecular formula of $C_{43}H_{62}O_{15}$. On the basis of its ^1H and ^{13}C NMR spectra, aglycone of compound **61** was identified as arjunolic acid (**58**). In addition, one aromatic singlet at δ_H 7.87 (2H, *s*, H-2′,6′) on its ^1H NMR spectrum and carbon signals at δ_C 167.6 (C-7′), 148.1 (C-3′,5′), 141.4 (C-4′), 122.0 (C-1′), and 110.5 (C-2′,6′) evidenced the presence of galloyl unit in compound **61** (Figure 98). The linkage position of the galloyl unit at C-23 position was confirmed from the cross peak correlation displayed on the HMBC spectrum between H-23 [δ_H 4.71 and 4.53 (each 1H, *d*, *J* = 11.1 Hz)] and C-7′. Furthermore, the presence of an anomeric proton signal at δ_H 6.33 (1H, *d*, *J* = 8.1 Hz, H-1″) and anomeric carbon signal at δ_C 96.2 (C-1″) confirmed the presence of a *β*-glucopyranose unit. The location of glucose moiety at C-28 position was determined by the HMBC cross peak correlation between the anomeric proton and C-28 (δ_C 176.9). Thus, compound **61** was elucidated as 23-*O*-galloylarjunolic acid 28-*O*-*β*-D-glucopyranosyl ester (Conrad et al., 1998).

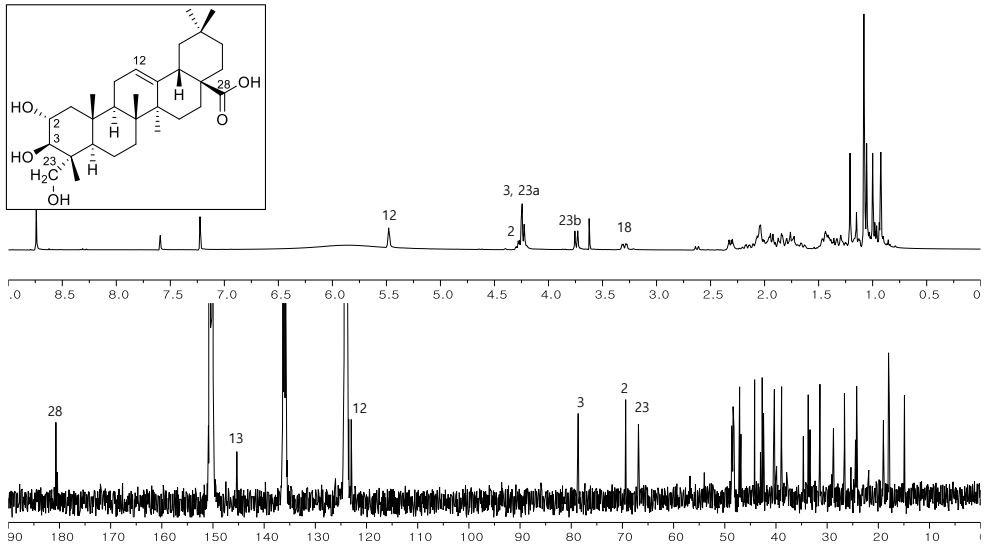


Figure 95. ^{1}H and ^{13}C NMR spectra of compound **58**

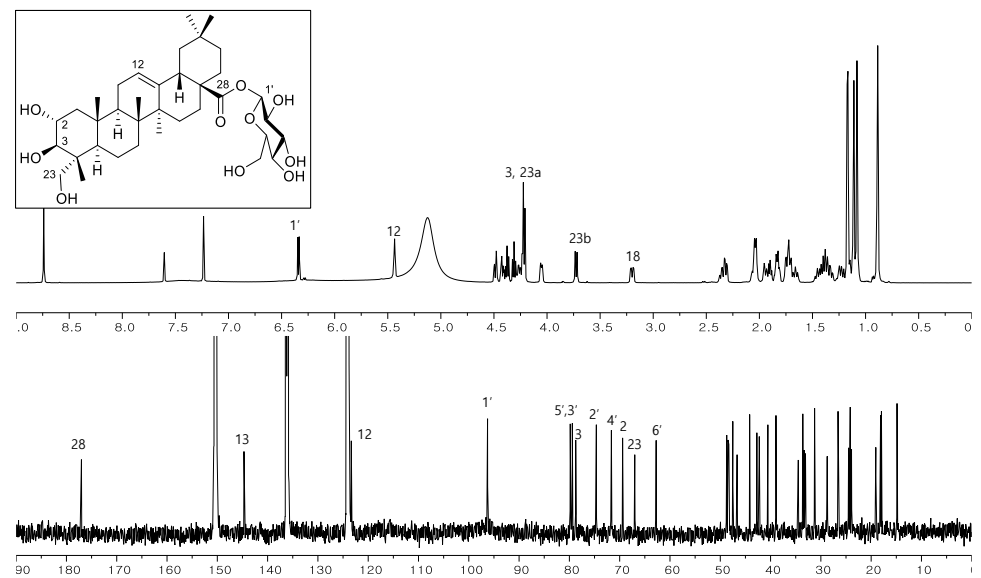


Figure 96. ^{1}H and ^{13}C NMR spectra of compound **59**

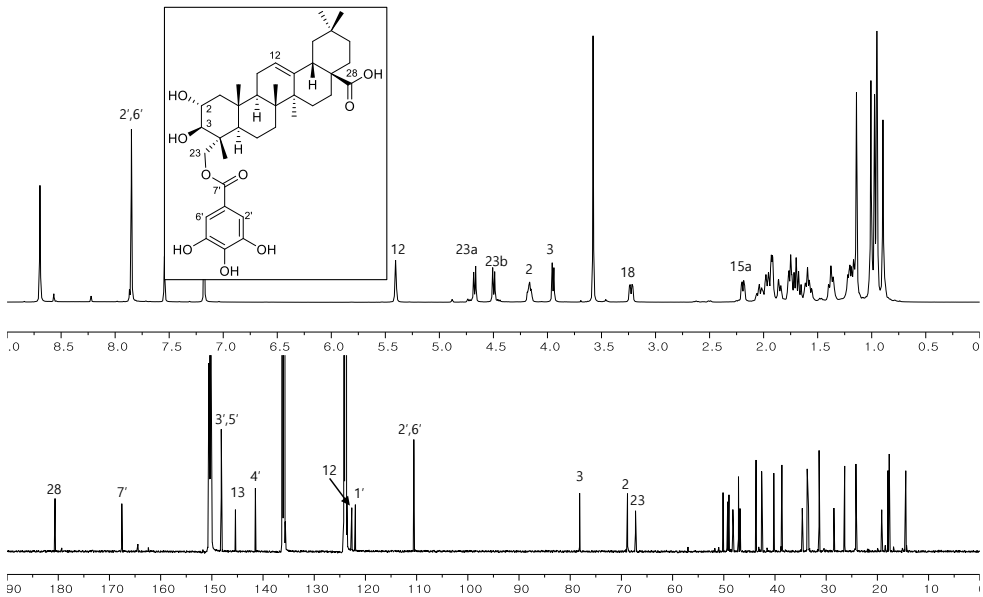


Figure 97. ^{1}H and ^{13}C NMR spectra of compound **60**

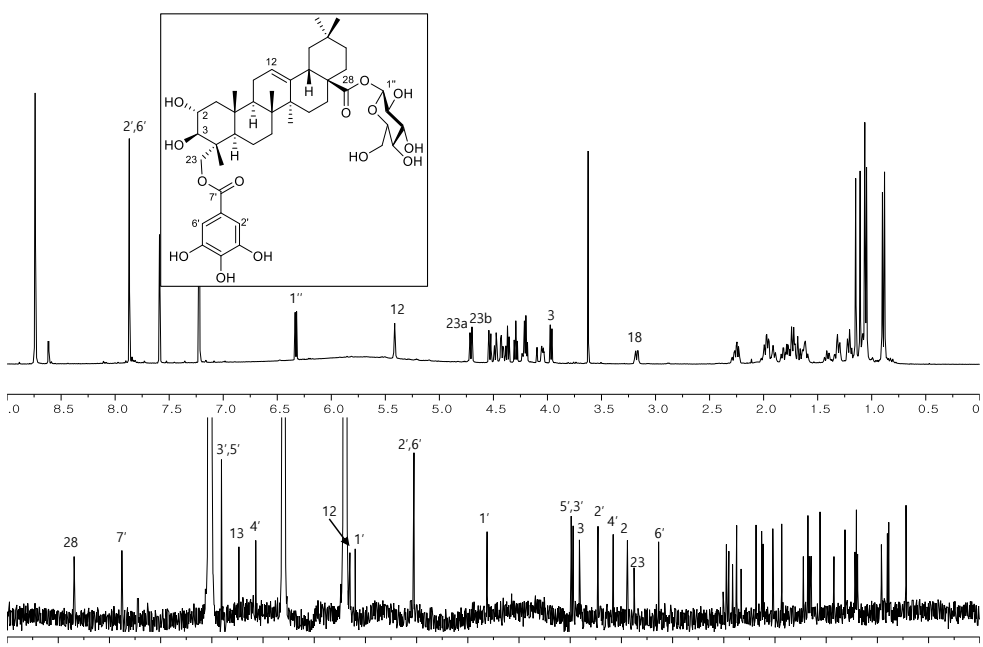


Figure 98. ^{1}H and ^{13}C NMR spectra of compound **61**

3.23. Compounds **62** and **63**

Compound **62** was obtained as white amorphous powder. The (-)-HRESIMS of **62** exhibited a quasi-molecular ion peak at *m*/*z*: 487.3432 $[M - H]^-$ indicating a molecular formula of $C_{30}H_{48}O_5$ (calcd for $C_{30}H_{47}O_5$ 487.3423). Its ^1H and ^{13}C NMR spectra (Figure 99) showed typical signals for olean-12-ene-28-oic acid skeleton including seven tertiary methyl groups [δ_H 1.61 (H-27), 1.27 (H-23), 1.17 (H-29), 1.10 (H-24), 1.10 (H-23), 1.06 (H-25), and 1.01 (H-26)] as well as an olefinic proton (δ_H 5.53, br *t*, *J* = 3.2 Hz, H-12), two olefinic carbons (δ_C 145.4 and 123.9), and a carboxyl carbon (δ_C 181.4). In addition, the 1H NMR spectrum of **62** exhibited three oxymethine protons at δ_H 4.08 (1H, *td*, *J* = 10.3, 4.3 Hz, H-2), 3.60 (1H, br *s*, H-19), and 3.36 (1H, *d*, *J* = 9.4 Hz, H-3), which indicated that compound **62** had three hydroxyl groups. The HMBC correlations indicated that four hydroxyl groups were located at C-2, C-3, and C-19. Coupling constant between H-2 and H-3 (*J* = 9.4 Hz) suggested that the hydroxyl groups at C-2 and C-3 were both equatorially oriented (2α, 3β). Based on these spectral data and comparison with literature, compound **62** was determined as arjunic acid (Joo et al., 2016).

Compound **63** was obtained as white amorphous powder and showed a molecular ion peak [M + HCOOH – H]- at *m*/*z* 695.4018 (calcd for $C_{37}H_{59}O_{12}$ 695.4007) in the negative HRESIMS consistent with the molecular formula $C_{36}H_{58}O_{10}$. Analysis of the ^1H and ^{13}C NMR spectra (Figure 100) of **63** indicated the presence of a β-glucopyranosyl moiety from signal at δ_H 6.41 (1H, *d*, *J* = 8.1 Hz, H-1') and an anomeric carbon at δ_C 95.7 (C-1'). Except for the signals corresponding to the β-glucopyranose moiety, the ^1H and ^{13}C NMR spectra of **57** was similar to those of **62**. The location of β-glucopyranose moiety of was determined to be the C-28 position by HMBC correlation between the signals of H-1′ (δ_H 6.41) and C-28 (δ_C 177.7). Thus compound **53** was confirmed as arjunetin (Zhou et al., 1992).

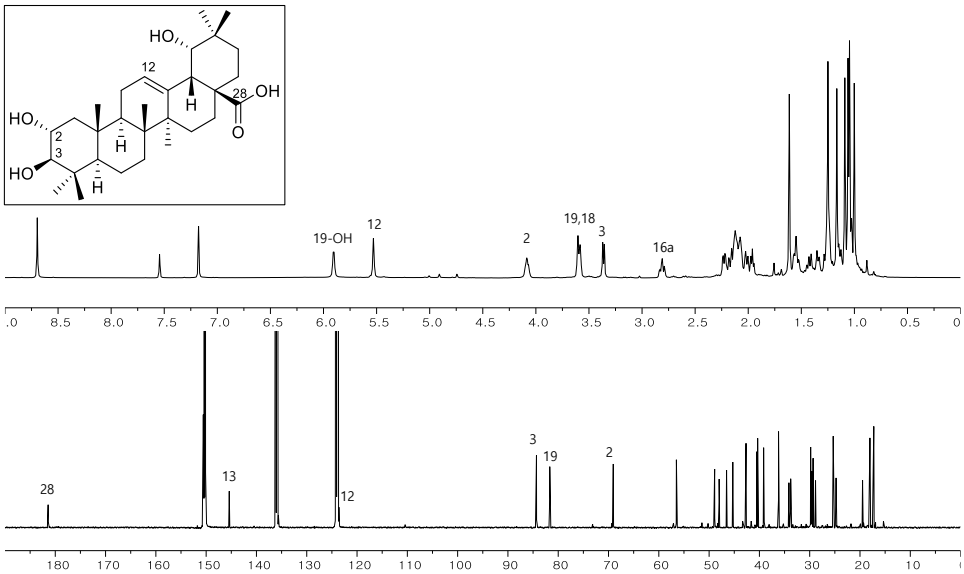


Figure 99. ^{1}H and ^{13}C NMR spectra of compound **62**

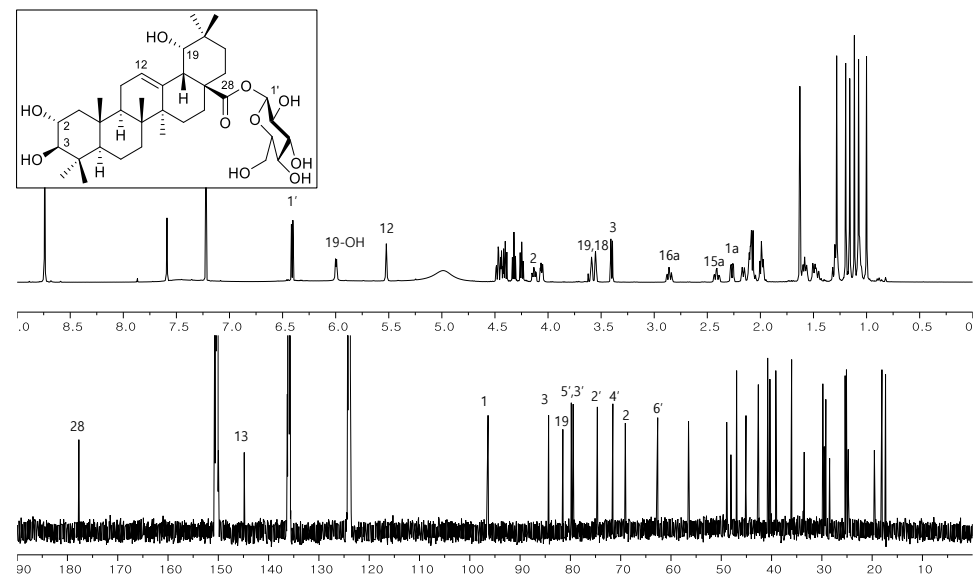


Figure 100. ^{1}H and ^{13}C NMR spectra of compound **63**

3.24. Compound **64**

Compound **64** was obtained as white amorphous powder Its molecular formula, $C_{36}H_{58}O_{10}$ was determined by (-)-HRESIMS (*m*/*z*: 695.4011 [M + HCOOH – H]$^{-}$,

calcd for $C_{37}H_{59}O_{12}$ 695.4007). The ^{1}H NMR spectrum (Figure 101) exhibited the presence of six methyls [δ_H 1.60 (H-27), 1.22 (H-26), 1.16 (H-29), 1.09 (H-24), 1.05 (H-25), and 1.00 (H-30)], one oxymethylene [δ_H 4.19 and 3.74 (each 1H, *d*, *J* = 10.3 Hz, H-23), two oxymethines [δ_H 4.23 (1H, *m*, H-3), 3.60 (1H, br *s*, H-19], and an olefinic protons [δ_H 5.55 (1H, br *t*, *J* = .4.0 Hz, H-12)]. These characters and the chemical shifts of one carboxyl carbon (δ_C 177.8) and two olefinic carbons (δ_C 144.8 and 124.3) suggested that compound **64** had olean-12-ene-28-oic acid skeleton. Furthermore, ^{1}H and ^{13}C NMR signals at δ_H 6.42 (1H, *d*, *J* = 8.1 Hz, H-1′) and six oxygenate protons between δ_H 4.47 and 4.06 ppm and δ_C 96.4 (C-1′), 79.8 (C-5′), 79.4 (C-3′), 74.6 (C-2′), 71.5 (C-4′) and 62.6 (C-6′) indicated the presence of β-glucopyranosyl moiety. The HMBC spectrum indicated that the three hydroxyl groups were existed at C-3, C-19, and C-23. The location of β-glucopyranose moiety of was also determined to be the C-28 position by HMBC correlation between the signals of H-1′ (δ_H 6.42) and C-28 (δ_C 177.8). On the basis of above spectral data and comparison with literature, compound **64** was identified as crataegioside (Jung et al., 2001).

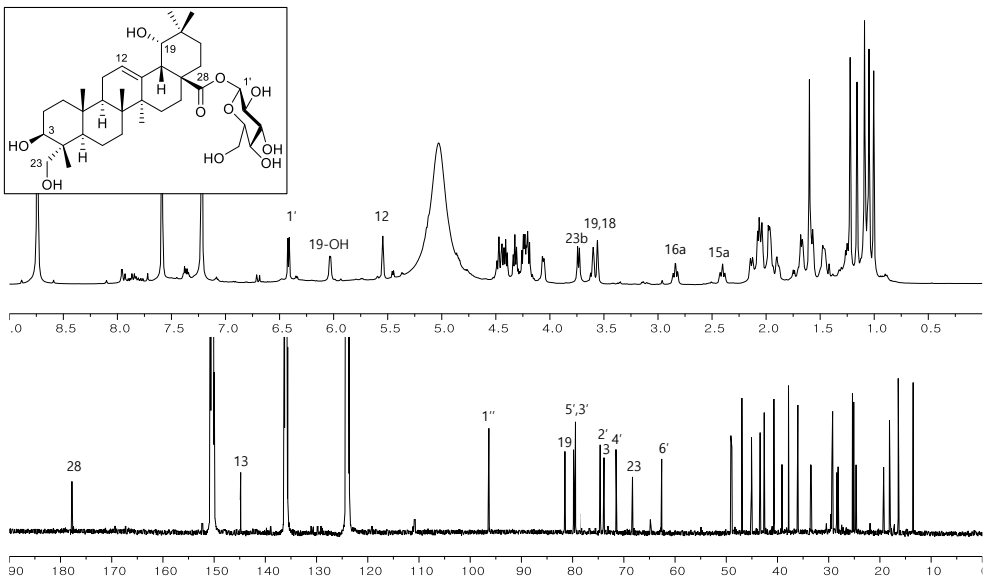


Figure 101. ^{1}H and ^{13}C NMR spectra of compound **64**

3.25. Compounds **65** and **66**

Compound **65**, white amorphous powder, was determined to have a molecular formula of $C_{36}H_{56}O_{10}$ by its (-)-HRESIMS peak at *m/z* 693.3864 [M + HCOOH – H]$^{-}$ (calcd for $C_{37}H_{57}O_{12}$ 693.3850). The ^{1}H NMR spectrum (Figure 102) showed five methyl singlets at δ_H 1.03 (H-27), 1.08 (H-24), 1.14 (H-25), 1.22 (H-26), and 1.79 (H-29), one methyl doublet at δ_H 1.02 (d, *J* = 7.0 Hz), two oxygen-bearing methine protons at δ_H 4.28 (1H, *m*, H-2) and 4.23 (1H, *d*, *J* = 9.3 Hz, H-3), oxygen-bearing methylene protons at δ_H 4.23 and 3.74 (each 1H, *d*, *J* = 10.4 Hz, H-23), and an olefinic proton at δ_H 5.71 (H-12, br *t*, *J* = 3.8 Hz). In addition, the ^{13}C NMR showed four olefinic signals (δ_C 127.1, 134.3, 136.4, and 139.3), of which δ_C 127.1 and 139.3 were typical of a double bond at C-12 and C-13 of an urs-12-ene-type triterpene, (Mahato and Kundu, 1994) and a carboxyl carbon signal at δ_C 175.2. The HMBC correlations indicated that a carboxyl group and three hydroxyl groups were existed at C-28, C-2, C-3, and C-23, respectively. The other double bond was also identified at C-18 – C-19 by the HMBC cross-peaks. Large coupling constant

of H-3 (J = 9.7 Hz) indicated that the hydroxyl moieties at C-2 and C-3 of **65** were oriented in the α-, β- position resepectively. Based on these data, the aglycone of **65** was identified as pinfaenoic acid (Durham et al., 1996). The presence of one glucopyranosyl moiety and its β-orientation were evident from the anomeric proton signal at δ_H 6.35 (H-1′, *d*, J = 8.2 Hz) and its typical signals at δ_C 96.3 (C-1′), 79.6 (C-5′), 79.4 (C-3′), 74.6 (C-2′), 71.7 (C-4′), and 62.8 (C-6′) and its location was determined to be the C-28 position by HMBC correlation between the signals of H-1′ (δ_H 6.35) and C-28 (δ_C 175.2). Based on these spectral data, compound **65** was identified as pinfaenoic acid 28-*O*-β-D-glycopyranosyl ester (Durham et al., 1996).

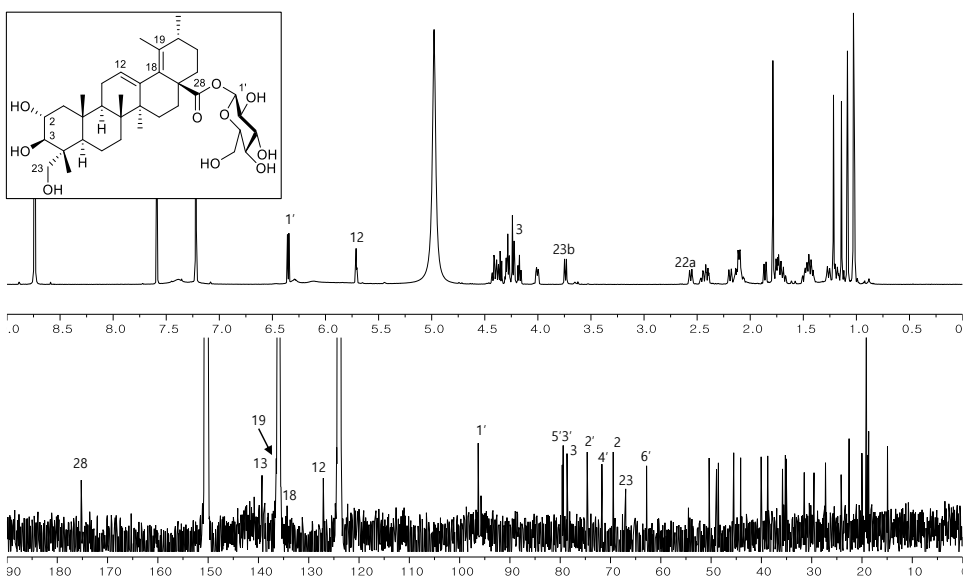


Figure 102. ^{1}H and ^{13}C NMR spectra of compound **65**

Compound **66** was obtained yellowish amorphous powder. Its molecular formula of $C_{43}H_{60}O_{14}$ was determined by the quasi-molecular ion peak at m/z 799.3889 [M – H]$^-$ (calcd for $C_{43}H_{59}O_{14}$ 693.3850). The ^{1}H and ^{13}C NMR of **66** were similar to those of compound **65**, except for the additional signals at δ_H 7.85 (2H, *s*, H-2′,6′) and δ_C 167.6 (C-7′), 148.2 (C-3′,5′), 141.4 (C-4′), 121.9 (C-1′), and 110.5 (C-2′,6′),

corresponding to galloyl unit, were displayed (Figure 103). The location of galloyl group at C-23 was indicated by the strongly downfield shift of H-23 (δ_H 4.75 and 4.50, each 1H, *d*, *J* = 11.0 Hz) and HMBC correlation between the signals of H-23 and C-7′ (Figure 104). Thus compound **66** was identified as 23-*O*-galloylpinfaenoic acid *β*-D-glycopyranosyl ester and it was isolated from nature for the first time.

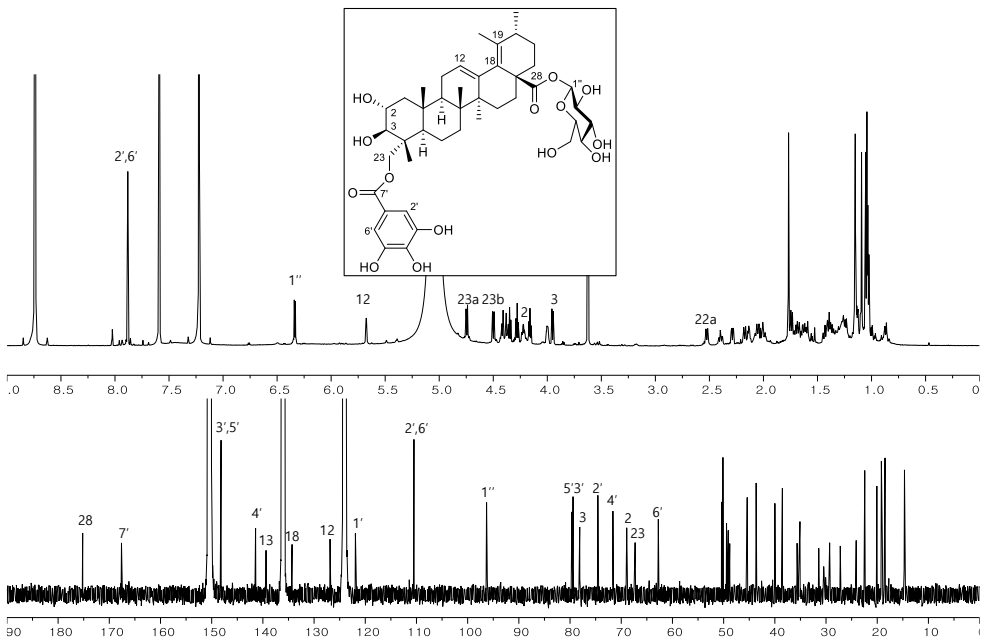


Figure 103. ^{1}H and ^{13}C NMR spectra of compound **66**

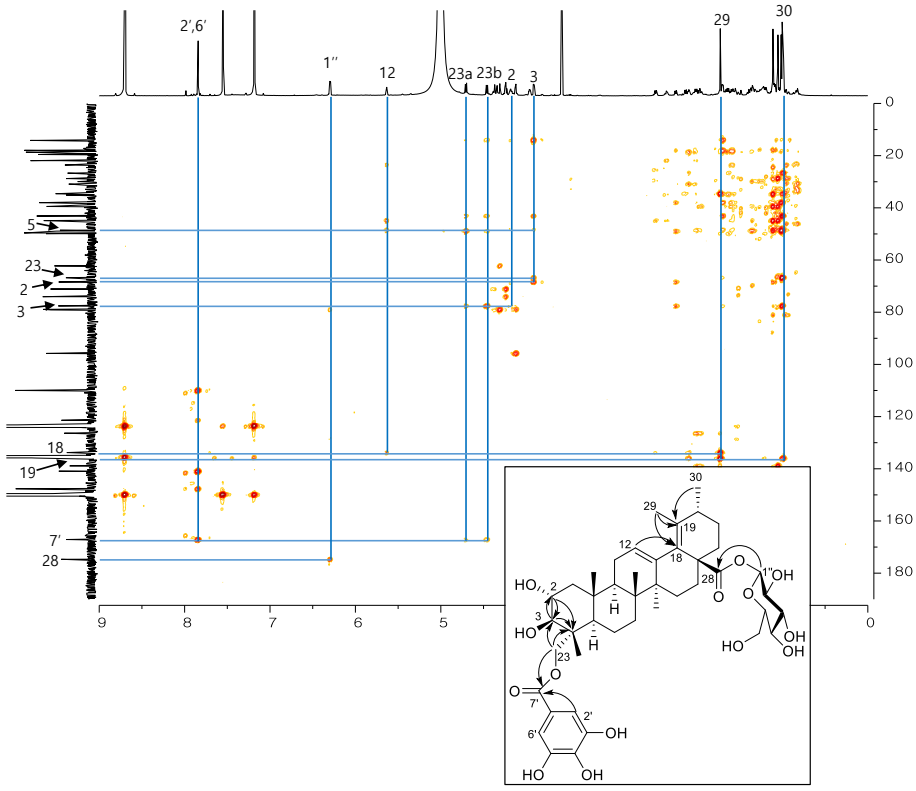
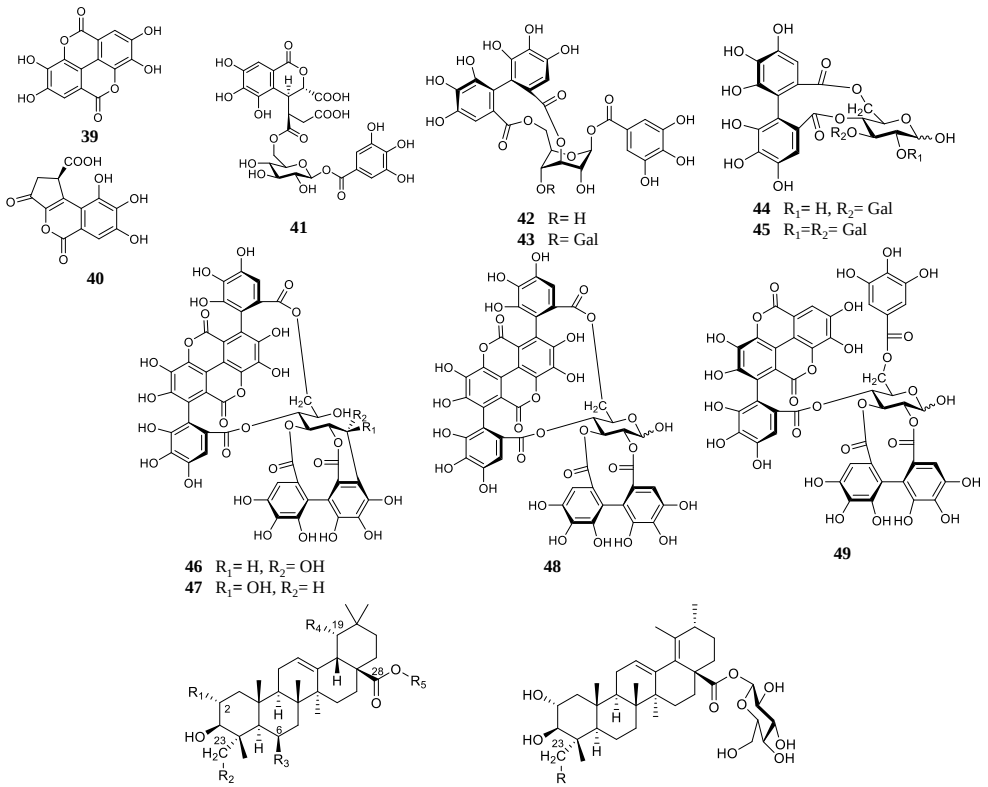


Figure 104. HMBC spectrum of compound **66**

Figure 105. Isolated compounds from *T. chebula* fruits (continued)



	R_1	R_2	R_3	R_4	R_5
50	OH	OH	H	OH	H
51	OH	N1	H	OH	*β*-D-glucose
52	OH	N2	H	OH	H
53	OH	*O*-Gal	H	OH	H
54	OH	*O*-Gal	H	OH	*β*-D-glucose
55	OH	OH	H	OH	*β*-D-glucose
56	OH	OH	OH	H	H
57	OH	*O*-Gal	OH	H	*β*-D-glucose
58	OH	OH	H	H	H
59	OH	OH	H	H	*β*-D-glucose
60	OH	*O*-Gal	H	H	H
61	OH	*O*-Gal	H	H	*β*-D-glucose
62	OH	H	H	OH	H
63	OH	H	H	OH	*β*-D-glucose
64	H	OH	H	OH	*β*-D-glucose

	R
65	OH
66	*O*-Gal

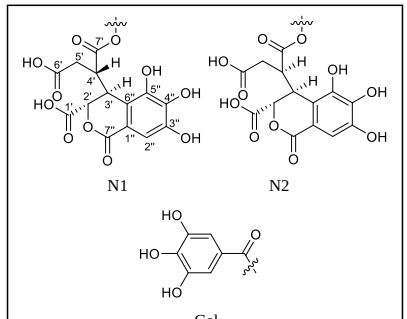


Figure 106. Isolated compounds from *T. chebula* fruits

4. Inhibitory activities of compounds from *T. chebula* against Baker's yeast α-glucosidase, rat intestinal α-glucosidase, porcine pancreatic α-amylase, and PTP-1B.

The IC_{50} values of 66 isolated compounds against Baker's yeast *α*-glucosidase were tested and are shown in Table 28. The results show remarkable variations in the inhibitory activity, due to their structural differences. Among these compounds, 4-*O*-(2″,4″-di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**11**, IC_{50} 6.4 μM) and 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**, IC_{50} 2.9 μM) showed most potent *α*-glucosidase inhibitory activity. The inhibitory activities of cinnamoyl-conatining gallotannins (IC_{50} 68.4–2.9 μM) (**12**, **13**, **16**, and **17**) were enhanced with the increase of a free galloyl unit. Rhamnosylated ellagic acid derivatives (**9**–**11**, and **26**), although eschweilenol C (**26**) had no inhibitory activity in this study, showed potent inhibitory activities (IC_{50} 37.9–6.4 μM) against *α*-glucosidase. In addition, they showed more potent inhibitory activity when compared with the simple gallotannins that have the same galloyl unit (**14**, **15**, **31**, and **32**). Hence, it was suggested that the 4-*O*-rhamnosyl ellagic acid group has a significant role in *α*-glucosidase inhibition. Among chebulic ellagitannins (**2**-**4**, **7**, **18**-**25**, and **41**) and chebulic acid derivatives (**34**–**37**), phyllanemblinins (**19**, **20**, and **41**), neochebulagic acid derivatives (**3**, **22**–**24**), and chebulic acid derivatives (**34**–**37**) had no inhibitory activities. Neochebulinic acid derivatives (**4** and **21**) and chebulinic acid (**7**), which have three galloyl units, showed moderate inhibitory activities (IC_{50} 59.5 25.4, and 24.0 μM, respectively). Chebulanin (**25**), which has a 2-4-*O*-chebuloyl unit and only one galloyl unit, had no inhibitory activity. These results suggested that the 4-*O*-neochebuloyl and 2-4-*O*-chebuloyl unit did not have a significant influence on *α*-glucosidase inhibition. Interestingly, chebulagic acid (**18**) had no inhibitory activity (IC_{50} >100 μM) whereas its methyl derivative methyl chebulagate (**18**) showed good inhibitory activity (IC_{50} 16.3 μM), which

indicated that the presence of methyl group in 2-4-*O*-chebuloyl unit of **18** played an important role in enzyme inhibition. Mono-, di-galloyl glucose (**14**, **15**, **31**, **32**), Simple gallic acid esters (**27**, **28**, and **38**), and monogalloyl shikimic acid (**29** and **30**) had no α-glucosidase inhibitory activities (IC_{50} >100 μM). Whereas, tetra- and penta- galloyl glucose (**5**, **6**, and **8**) showed significant inhibitory activities (IC_{50} 15.7–8.3 μM). When compared the inhibition activity of **42** (IC_{50} >200 μM), and **43** (IC_{50} 48.1 μM) with their respective structural isomer **44** (IC_{50} 62.7 μM), and **45** (IC_{50} 29.3 μM), 4-6-*O*-hexahydroxyphenyl (HHDP) moiety played more important role than 3-6-*O*-HHDP moiety in α-glucosidase inhibition. Compound **46**–**49**, with bulky a gallagyl (**46**–**48**) or flavogallonyl unit (**49**), showed moderate inhibitory activity (IC_{50} 38.3-18.7 μM). Compounds **60** (IC_{50} 21.7 μM) and **61** (IC_{50} 64.2 μM), which have the same aglycone, arjunolic acid, and 23-*O*-galloyl moiety were most active compounds among triterpene derivatives. Other triterpene derivatives (**50**-**59** and **62**-**66**) did not show inhibitory activity.

Table 28. Baker's yeast α-glucosidase inhibitory activity of compounds **1**–**66**

Compound	IC_{50} (μM)	Compound	IC_{50} (μM)	Compound	IC_{50} (μM)
1	> 100	**23**	> 100	**45**	29.3 ± 1.2
2	> 100	**24**	> 100	**46**	38.3 ± 6.5
3	> 100	**25**	24.0 ± 4.0	**47**	35.5 ± 2.4
4	59.5 ± 9.6	**26**	> 100	**48**	18.7 ± 3.3
5	15.7 ± 0.8	**27**	> 100	**49**	19.8 ± 1.3
6	15.5 ± 0.6	**28**	> 100	**50**	> 100
7	24.0 ± 4.0	**29**	> 100	**51**	> 100
8	8.3 ± 0.5	**30**	> 100	**52**	> 100
9	37.9 ± 6.7	**31**	> 100	**53**	> 100
10	12.3 ± 2.0	**32**	> 100	**54**	> 100
11	6.4 ± 1.0	**33**	> 100	**55**	> 100
12	2.9 ± 0.8	**34**	> 100	**56**	> 100
13	14.7 ± 2.7	**35**	> 100	**57**	> 100
14	> 100	**36**	> 100	**58**	> 100

15	> 100	**37**	> 100	**59**	> 100
16	46.1 ± 5.1	**38**	> 100	**60**	21.7 ± 5.6
17	68.4 ± 7.5	**39**	> 100	**61**	64.2 ± 13.6
18	16.3 ± 2.1	**40**	> 100	**62**	> 100
19	> 100	**41**	> 100	**63**	> 100
20	> 100	**42**	> 100	**64**	> 100
21	> 100	**43**	48.1 ± 7.1	**65**	> 100
22	25.4 ± 2.4	**44**	62.7 ± 5.6	**66**	> 100

Acarbose (positive control): 174.0 ± 29.8 μM

*All the assays were conducted in triplicate.

*IC_{50} values represent means ± standard deviation.

To determine the mode of inhibition of Baker's yeast *α*-glucosidase, 4-*O*-(2″,4″-di-*O*-galloyl-α-L-rhamnosyl) ellagic acid (**11**) and 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**), which were the most potent inhibitors in this study, were analyzed by Michaelis-Menten studies. In this study, as shown in Figure 107, the Lineweaver-Burk plots of **11** and **12** had no intersection on the x or y axis, indicating that the mode of inhibition was the mixed-inhibition mode. Furthermore, Ki, the inhibition constants, values were 4.0 and 1.9 μM, respectively. Acarbose, the positive control used in this study, displayed competitive inhibition against *α*-glucosidase. The Inhibition mode of acarbose in this study was consistent with previous studies (Kashtoh et al., 2014; Sivasothy et al., 2016).

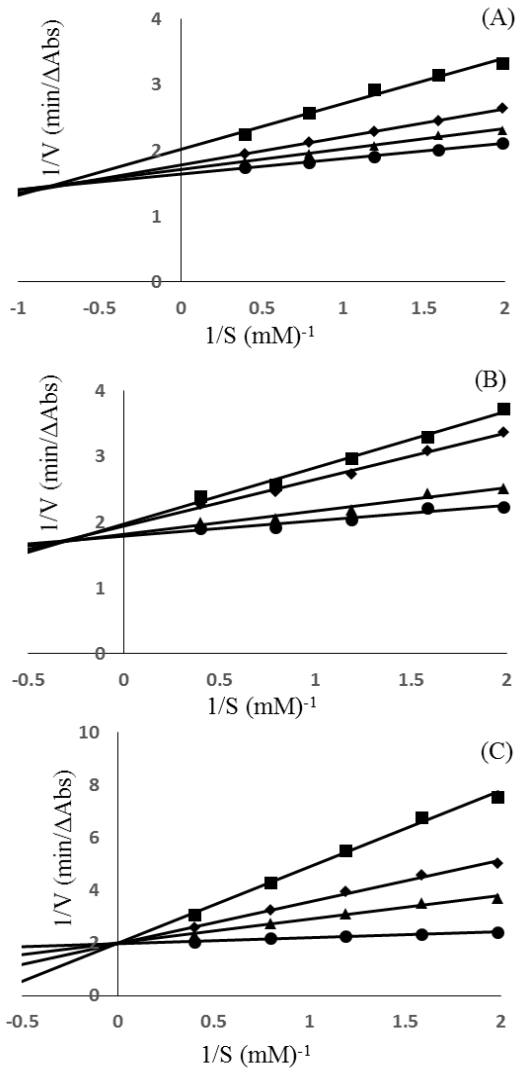


Figure 107. Lineweaver-Burk double-reciprocal plots for the inhibition of Baker's yeast α-glucosidase in the presence of (A) 4-*O*-(2″,4″-di-*O*-galloyl-α-l-rhamnosyl)ellagic acid (**11**), (B) 1,3,4,6-tetra-*O*-galloyl-2-*O*-cinnamoyl-*β*-d-glucose (**12**), and (C) acarbose (positive control).

Table 29 and 30 showed the inhibitory activity against rat intestinal α-glucosidase and porcine pancreatic α-amylase of all isolated compounds.

Table 29. Rat intestinal α-glucosidase inhibition of compounds **1**-**66**

Cpd. No.	Inhibition (%)[a]	Cpd. No.	Inhibition (%)[a]	Cpd. No.	Inhibition (%)[a]
1	33.1 ± 3.6	**23**	12.9 ± 0.6	**45**	36.5 ± 5.5
2	11.3 ± 0.1	**24**	7.9 ± 0.4	**46**	9.2 ± 0.6
3	3.1 ± 0.2	**25**	16.5 ± 0.1	**47**	6.2 ± 0.2
4	18.0 ± 0.7	**26**	4.9 ± 0.1	**48**	25.0 ± 6.0
5	47.7 ± 8.7	**27**	NI[b]	**49**	31.7 ± 4.4
6	56.4 ± 8.8	**28**	2.7 ± 0.2	**50**	1.4 ± 0.3
7	22.3 ± 0.2	**29**	NI[b]	**51**	3.4 ± 0.7
8	98.1 ± 6.3	**30**	0.8 ± 0.1	**52**	4.3 ± 0.3
9	4.3 ± 0.2	**31**	3.1 ± 0.1	**53**	7.4 ± 0.2
10	8.6 ± 0.1	**32**	10.1 ± 0.8	**54**	4.1 ± 0.2
11	7.7 ± 0.3	**33**	28.7 ± 1.2	**55**	NI[b]
12	21.4 ± 0.8	**34**	22.1 ± 2.0	**56**	5.1 ± 0.6
13	10.4 ± 0.8	**35**	24.7 ± 3.4	**57**	4.2 ± 0.7
14	3.1 ± 0.2	**36**	17.0 ± 2.4	**58**	2.4 ± 0.6
15	10.3 ± 0.7	**37**	21.3 ± 2.6	**59**	3.2 ± 1.2
16	6.6 ± 0.4	**38**	2.1 ± 0.1	**60**	22.1 ± 1.8
17	7.1 ± 0.5	**39**	29.5 ± 1.6	**61**	14.3 ± 0.5
18	20.3 ± 0.1	**40**	12.1 ± 0.9	**62**	3.7 ± 0.5
19	18.9 ± 0.2	**41**	2.5 ± 0.1	**63**	3.8 ± 1.4
20	20.4 ± 1.1	**42**	9.8 ± 0.3	**64**	4.8 ± 0.5
21	23.8 ± 1.4	**43**	30.2 ± 2.8	**65**	3.6 ± 0.2
22	3.4 ± 0.2	**44**	23.9 ± 4.5	**66**	2.5 ± 0.3
Acarbose[c]	58.6 ± 4.9				

[a] Percentage of inhibition at the concentration of 100 μM. Data presented is the mean ± SD of sample runs in triplicate.

[c] No inhibition.

[d] Positive control

Table 30. Porcine pancreatic α-amylase inhibition of compounds **1-66**

Cpd. No.	Inhibition (%)[a]	Cpd. No.	Inhibition (%)[a]	Cpd. No.	Inhibition (%)[a]
1	6.6 ± 0.2	**23**	4.7 ± 0.2	**45**	0.5 ± 0.1
2	1.3 ± 0.1	**24**	5.6 ± 0.1	**46**	1.7 ± 0.1
3	5.0 ± 0.1	**25**	1.1 ± 0.1	**47**	1.1 ± 0.1
4	2.3 ± 0.1	**26**	1.6 ± 0.2	**48**	1.4 ± 0.1
5	8.4 ± 0.5	**27**	2.8 ± 0.1	**49**	1.7 ± 0.1
6	7.6 ± 0.2	**28**	4.0 ± 0.1	**50**	0.8 ± 0.1
7	1.7 ± 0.1	**29**	4.9 ± 0.1	**51**	2.2 ± 0.2
8	94.2 ± 0.9	**30**	6.2 ± 0.1	**52**	1.9 ± 0.1
9	NI[b]	**31**	1.9 ± 0.1	**53**	0.8 ± 0.2
10	2.7 ± 0.2	**32**	5.8 ± 0.1	**54**	NI[b]
11	2.9 ± 0.1	**33**	7.8 ± 0.4	**55**	NI[b]
12	11.0 ± 1.1	**34**	4.2 ± 0.2	**56**	3.8 ± 0.4
13	7.0 ± 0.3	**35**	5.3 ± 0.3	**57**	1.1 ± 0.2
14	4.8 ± 0.3	**36**	4.2 ± 0.1	**58**	4.9 ± 0.2
15	4.7 ± 0.2	**37**	4.9 ± 0.2	**59**	2.9 ± 0.3
16	4.5 ± 0.1	**38**	7.6 ± 0.1	**60**	4.3 ± 0.7
17	5.7 ± 0.2	**39**	1.2 ± 0.1	**61**	2.4 ± 0.2
18	NI[b]	**40**	4.8 ± 0.1	**62**	3.1 ± 0.4
19	5.9 ± 0.1	**41**	2.5 ± 0.1	**63**	3.6 ± 0.2
20	6.2 ± 0.1	**42**	1.0 ± 0.1	**64**	2.8 ± 0.3
21	4.0 ± 0.1	**43**	0.9 ± 0.1	**65**	1.2 ± 0.1
22	4.3 ± 0.1	**44**	0.8 ± 0.1	**66**	1.7 ± 0.2
Acarbose[c]	72.1 ± 1.5				

[a] Percentage of inhibition at the concentration of 100 μM. Data presented is the mean ± SD of sample runs in triplicate.

[c] No inhibition.

[d] Positive control

In the case of rat intestinal alpha glucosidase and porcine pancreatic alpha amylase, only compound **8** showed potent inhibitory activities compared to the positive control, acarbose. The IC_{50} values of **8** against rat intestinal alpha glucosidase and porcine pancreatic alpha amylase were 17.3 ± 2.2 and 13.4 ± 0.8 μM, respectively. Acarbose, the positive control used in this study, showed rat intestinal alpha glucosidase and porcine pancreatic alpha amylase inhibitory activities with IC_{50} values 30.0 ± 3.6 and 27.3 ± 2.0 μM, respectively, which was similar to previous reports (Boath et al., 2012; Ponnusamy et al., 2012).

The inhibitory effects of all isolated compounds against PTP-1B were tested and are shown in Table 31.

Table 31. PTP-1B inhibitory activity of compounds **1**–**66**

Compound	IC_{50} (μM)	Compound	IC_{50} (μM)	Compound	IC_{50} (μM)
1	> 50	**23**	> 50	**45**	> 50
2	> 50	**24**	> 50	**46**	26.6 ± 8.5
3	> 50	**25**	> 50	**47**	20.5 ± 4.9
4	> 50	**26**	> 50	**48**	9.4 ± 1.4
5	7.7 ± 1.2	**27**	> 50	**49**	18.2 ± 3.7
6	10.2 ± 2.3	**28**	> 50	**50**	> 50
7	20.3 ± 3.8	**29**	> 50	**51**	> 50
8	2.0 ± 0.4	**30**	> 50	**52**	44.5 ± 9.7
9	25.2 ± 4.3	**31**	> 50	**53**	25.2 ± 5.6
10	17.2 ± 4.2	**32**	> 50	**54**	> 50
11	7.2 ± 1.5	**33**	> 50	**55**	> 50
12	11.4 ± 0.9	**34**	> 50	**56**	19.8 ± 3.1
13	11.9 ± 3.8	**35**	> 50	**57**	> 50
14	> 50	**36**	> 50	**58**	12.3 ± 2.7
15	> 50	**37**	> 50	**59**	> 50
16	> 50	**38**	> 50	**60**	10.3 ± 2.5
17	> 50	**39**	> 50	**61**	> 50
18	> 50	**40**	> 50	**62**	36.0 ± 6.5
19	> 50	**41**	> 50	**63**	> 50

20	> 50	**42**	> 50	**64**	46.3 ± 9.5
21	> 50	**43**	> 50	**65**	> 50
22	> 50	**44**	> 50	**66**	> 50

Oleanolic acid (positive control): 12.0 ± 3.4 μM

*All the assays were conducted in triplicate.

*IC_{50} values represent means ± standard deviation.

Among hydrolysable tannins (**1**-**49**), 1,2,3,4,6-penta-*O*-galloyl-*β*-D-glucose (**8**, IC_{50} 2.0 μM) showed most potent inhibitory activity. These results showed that the inhibitory activities of ellagic acid rhamnosides (**9**–**11**, and **26**) and cinnamoyl-containing gallotannins (**12**, **13**, **16**, and **17**) were enhanced with the increase of a free galloyl unit. Mono-, di-, and tri- galloyl glucose (**1**, **14**, **15**, and **31**-**33**) did not exhibit inhibitory activity, but tetra- and penta- galloyl glucose (**5**, **6**, and **8**) showed significant inhibitory activities (IC_{50} 10.2–2.0 μM). Simple gallic acid esters (**27**, **28**, and **38**), monogalloyl shikimic acid (**29** and **30**), and simple phenolic acid (**34**-**37**, **39**, and **40**) had no inhibitory activities (IC_{50} >50 μM). Except compound **7**, all of the ellagitannins (**42**-**45**) and the chebulic ellagitannins (**3**, **4**, **18**-**25**, and **41**) did not exhibit inhibitory activity. Compound **46**–**49**, with bulky a gallagyl (**46**–**48**) or flavogallonyl unit (**49**), showed inhibitory activity (IC_{50} 26.6-9.4 μM).

The inhibitory activity against PTP-1B of seventeen triterpene derivatives (**50**-**66**) were also evaluated. Among them, some of triterpene derivatives (**52**, **53**, **56**, **58**, **60**, **62**, and **64**) showed inhibitory activities against PTP-1B. Generally, triterpenes which have 28-*O*-*β*-D-glucopyranosyl ester unit (**51**, **54**, **55**, **57**, **59**, **63**, **65**, and **66**) showed no inhibitory activity (IC_{50} > 50 μM). Except compound **52** and **53**, arjungenin (2*α*,3*β*,19*α*,23-tetrahydroxyolean-12-en-28-oic acid) type compounds showed no inhibitory activity. Compound **52** and **53**, 23-*O*-4′-epi-neochebuloylarjungenin and 23-*O*-galloylarjunic acid, respectively, exhibited

moderate inhibitory activity (IC_{50} 44.5 and 25.2 μM, respectively). Three oleanolic acid type triterpenes (**56**, **62**, and **64**) also showed moderate inhibitory activity (IC_{50} 19.8, 36.0, and 46.3 μM, respectively). Especially, arjunolic acid (2*α*,3*β*,23-trihydroxyolean-12-en-28-oic acid) and 23-*O*-galloylarjunolic acid showed high inhibitory activity (IC_{50} 12.3 and 10.3 μM, respectively). Oleanolic acid, the positive control used in this study, showed PTP-1B inhibitory activities with IC_{50} values 12.0 ± 3.4 μM, which was analogous to previous studies (Ramirez-Espinosa et al., 2014).

IV. Conclusion

1. In order to isolate the new inhibitory compounds against α-glucosidase efficiently, HPLC-based activity profiling with dereplication approach was used.

2. Targeted isolation of thirteen hydrolysable tannins (**1**-**13**) from *T. chebula* fruits was performed by HPLC-based activity profiling. Among them, four compounds, 6′-*O*-methyl neochebulagate (**3**), 4-*O*-(2″,4″-di-*O*-galloyl-α-L-rhamnosyl)ellagic acid (**11**), 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**), and 1,2,3-tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (**13**), were reported for the first time from nature.

3. Evaluation of inhibitory activity of thirteen compounds against Baker's yeast *α*-glucosidase showed that compound **4**-**13** which was from active microfractions had inhibitory activities. Based on these results, it is confirmed that the HPLC-based activity profiling is an effective method for tracking targeted bioactive compounds.

4. In addition, in order to investigate on constituents of *T. chebula*, further isolation was performed. As a results, fifty-three compounds including thirty-six hydrolysable tannins (**14**-**49**) and seventeen polyhydroxytriterpenes (**50**-**66**) were isolated from MeOH extract of *T. chebula* fruits. Among them, eight compounds, methyl chebulagate (**18**), 1′-*O*-methyl neochebulanin (**20**), dimethyl neochebulinate (**21**), dimethyl neochebulagate (**23**), dimethyl 4′-*epi*-neochebulagate (**24**), 23-*O*-neochebuloylarjungenin 28-*O*-*β*-D-glucopyranosyl ester (**51**), 23-*O*-4′-*epi*-neochebuloylarjungenin (**52**), and 23-*O*-galloylpinfaenoic acid 28-*O*-*β*-D-glucopyranosyl ester (**66**) were newly reported from nature.

5. Structure elucidation of twelve new compounds was performed by spectroscopic data including NMR, HRMS, UV, and CD spectra and hydrolysis.

6. Among these isolated compounds, twenty-four compounds (**4**-**13**, **16**-**18**, **22**, **25**, **43**-**49**, **60**, and **61**) showed Baker's yeast α-glucosidase inhibitory activity. Especially, Compound **11** and **12** showed most potent inhibitory activity.

7. Only compound **8** exhibited potent inhibitory activity against rat intestinal α-glucosidase and porcine pancreatic α-amylase compared to positive control, acarbose.

8. In the case of PTP-1B, twenty compounds (**5**-**13**, **46**-**49**, **52**, **53**, **56**, **58**, **60**, **62**, and **64**) showed inhibitory activity. Among hydrolysable tannins, compound **8** showed the most potent inhibitory activity. And compound **60** was the most active compounds among triterpenes.

V. References

Abdel-Mageed, W. M., Bayoumi, S. A. H., Chen, C. X., Vavricka, C. J., Li, L., Malik, A., Dai, H. Q., Song, F. H., Wang, L. Q., Zhang, J. Y., Gao, G. F., Lv, Y. L., Liu, L. H., Liu, X. T., Sayed, H. M., Zhang, L. X., 2014. Benzophenone C-glucosides and gallotannins from mango tree stem bark with broad-spectrum anti-viral activity. *Bioorgan Med Chem* 22, 2236-2243.

Abdjul, D. B., Yamazaki, H., Takahashi, O., Kirikoshi, R., Ukai, K., Namikoshi, M., 2016. Sesquiterpene hydroquinones with protein tyrosine phosphatase 1B inhibitory activities from a dysidea sp. marine sponge collected in Okinawa. *J Nat Prod* 79, 1842-1847.

Abe, F., Yamauchi, T., 1987. Glycosides of 19-α-hydroxyoleanane-type triterpenoids from *Trachelospermum asiaticum*. *Chem Pharm Bull* 35, 1833-1838.

Agrawal, P. K., 1992. NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. *Phytochemistry* 31, 3307-3330.

Ahn, M. J., Kim, C. Y., Lee, J. S., Kim, T. G., Kim, S. H., Lee, C. K., Lee, B. B., Shin, C. G., Huh, H., Kim, J., 2002. Inhibition of HIV-1 integrase by galloyl glucoses from *Terminalia chebula* and flavonol glycoside gallates from *Euphorbia pekinensis*. *Planta Medica* 68, 457-459.

Anwesa, B., Subir, K. B., Rabi, R. C., 2013. The development of *Terminalia chebula* Retz. (Combretaceae) in clinical research. *Asian Pac J Trop Biomed* 3, 244-252.

Babu, K. S., Tiwari, A. K., Srinivas, P. V., All, A. Z., Raju, B. C., Rao, J. M., 2004. Yeast and mammalian α-glucosidase inhibitory constituents from Himalayan rhubarb *Rheum emodi* Wall.ex Meisson. *Bioorg Med Chem Lett* 14, 3841-3845.

Barthakur, N. N., Arnold, N. P., 1991. Nutritive value of the chebulic myrobalan (*Terminalia chebula* Retz) and its potential as a food source. *Food Chem* 40, 213-219.

Bisoli, E., Garcez, W. S., Hamerski, L., Tieppo, C., Garcez, F. R., 2008. Bioactive pentacyclic triterpenes from the Stems of *Combretum laxum*. *Molecules* 13, 2717-2728.

Boath, A. S., Stewart, D., McDougall, G. J., 2012. Berry components inhibit α-glucosidase in vitro: Synergies between acarbose and polyphenols from black currant and rowanberry. *Food Chem* 135, 929-936.

Chang, C. L., Lin, C. S., 2012. Phytochemical composition, antioxidant activity, and neuroprotective effect of *Terminalia chebula* Retz. extracts. *Evid Based Compl* Alt, 1-7.

Cheng, H. Y., Lin, T. C., Yu, K. H., Yang, C. M., Lin, C. C., 2003. Antioxidant and free radical scavenging activities of *Terminalia chebula*. *Biol Pharm Bull* 26, 1331-1335.

Chiasson, J. L., Josse, R. G., Gomis, R., Hanefeld, M., Karasik, A., Laakso, M., Group, S.-N. T. R., 2002. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet* 359, 2072-2077.

Conrad, J., Vogler, B., Klaiber, I., Roos, G., Walter, U., Kraus, W., 1998. Two triterpene esters from *Terminalia macroptera* bark. *Phytochemistry* 48, 647-650.

Cragg, G. M., Newman, D. J., 2013. Natural products: A continuing source of novel drug leads. *Bba-Gen Subjects* 1830, 3670-3695.

Debnath, J., Sharma, U. R., Kumar, B., Chauhan, N. S., 2010. Anticonvulsant activity of ethanolic extract of fruits of *Terminalia chebula* on experimental

animals. *Int J Drug Dev Res* 2, 764-768.

Ding, G., Liu, Y. Z., Lu, Y. R., Sheng, L. S., 2000. Two new isomers of chebulic acid from *Terminalia chebula*. *J Chin Pharm Univ* 32, 333-335.

Durham, D. G., Liu, I. X., Richards, R. M. E., 1996. Unsaturated E-ring triterpenes from *Rubus pinfaensis*. *Phytochemistry* 42, 505-508.

El-Toumy, S. A. A., Rauwald, H. W., 2002. Two ellagitannins from *Punica granatum* heartwood. *Phytochemistry* 61, 971-974.

Era, M., Matsuo, Y., Shii, T., Saito, Y., Tanaka, T., Jiang, Z. H., 2015. Diastereomeric ellagitannin isomers from *Penthorum chinense*. *J Nat Prod* 78, 2104-2109.

Escandon-Rivera, S., Gonzalez-Andrade, M., Bye, R., Linares, E., Navarrete, A., Mata, R., 2012. *α*-glucosidase inhibitors from *Brickellia cavanillesii*. *J Nat Prod* 75, 968-974.

Fahmy, N. M., Al-Sayed, E., Singab, A. N., 2015. Genus *Terminalia*: A phytochemical and biological review. *Med Aromat Plants* 4, 218-238.

Feldman, K. S., Sambandam, A., 1995. Ellagitannin chemistry. The first total chemical synthesis of an O(2),O(3)-galloyl-coupled ellagitannin, sanguiin H-5. *J Org Chem* 60, 8171-8178.

Foo, L. Y., 1993. Amariin, a di-dehydrohexahydroxydiphenoyl hydrolyzable tannin from *Phyllanthus amarus*. *Phytochemistry* 33, 487-491.

Goriparti, S., Harish, M. N. K., Sampath, S., 2013. Ellagic acid - a novel organic electrode material for high capacity lithium ion batteries. *Chem Commun* 49, 7234-7236.

Gossan, D. P. A., Magid, A. A., Yao-Kouassi, P. A., Josse, J., Gangloff, S. C., Morjani, H., Voutquenne-Nazabadioko, L., 2016. Antibacterial and cytotoxic triterpenoids from the roots of *Combretum racemosum*. *Fitoterapia* 110, 89-95.

Gu, D. Y., Yang, Y., Bakri, M., Chen, Q. B., Xin, X. L., Aisa, H. A., 2013. A LC/QTOF-MS/MS Application to investigate chemical compositions in a fraction with protein tyrosine phosphatase 1B inhibitory activity from *Rosa rugosa* flowers. *Phytochem Ana* 24, 661-670.

Guo, S., 2014. Insulin signaling, resistance, and the metabolic syndrome: insights from mouse models into disease mechanisms. *J Endocrinol* 220, T1-T23.

Haddock, E. A., Gupta, R. K., Alshafi, S. M. K., Layden, K., Haslam, E., Magnolato, D., 1982. The metabolism of gallic acid and hexahydroxydiphenic acid in plants. 4. Biogenetic and molecular taxonomic considerations. *Phytochemistry* 21, 1049-1062.

Hagenah, S., Gross, G. G., 1993. Biosynthesis of 1,2,3,6-tetra-*O*-galloyl-*β*-D-glucose. *Phytochemistry* 32, 637-641.

Hamada, S., Kataoka, T., Woo, J. T., Yamada, A., Yoshida, T., Nishimura, T., Otake, N., Nagai, K., 1997. Immunosuppressive effects of gallic acid and chebulagic acid on CTL-mediated cytotoxicity. *Biol Pharm Bull* 20, 1017-1019.

Hanhineva, K., Rogachev, I., Kokko, H., Mintz-Oron, S., Venger, I., Karenlampi, S., Aharoni, A., 2008. Non-targeted analysis of spatial metabolite composition in strawberry (*Fragariax ananassa*) flowers. *Phytochemistry* 69, 2463-2481.

Hasani-Ranjbar, S., Vahidi, H., Taslimi, S., Karimi, N., Larijani, B., Abdollahi, M., 2010. A systematic review on the efficacy of herbal medicines in the management of human drug-induced hyperprolactinemia; Potential sources for the development

of novel drugs. *Int J Pharmacol* 6, 691-695.

Hatano, T., Yoshida, T., Shingu, T., Okuda, T., 1988. C-13 Nuclear magnetic resonance spectra of hydrolyzable tannins. 2. Tannins forming anomer mixtures. *Chem Pharm Bull* 36, 2925-2933.

Hirsh, A. J., Yao, S. Y. M., Young, J. D., Cheeseman, C. I., 1997. Inhibition of glucose absorption in the rat jejunum: A novel action of *α*-D-glucosidase inhibitors. *Gastroenterology* 113, 205-211.

Hostettmann, K., Marston, A., Wolfender, J. L., 2005. Strategy in the search for new lead compounds and drugs from plants. *Chimia* 59, 291-294.

Hsu, F. L., Lee, Y. Y., Cheng, J. T., 1994. Antihypertensive activity of 6-*O*-galloyl-D-glucose a phenolic glycoside from *Sapium sebiferum*. *J Nat Prod* 57, 308-312.

Hua, J., Qi, J., Yu, B. Y., 2014. Iridoid and phenylpropanoid glycosides from *Scrophularia ningpoensis* Hemsl. and their *α*-glucosidase inhibitory activities. *Fitoterapia* 93, 67-73.

Huang, J., Wang, Y. H., Li, C., Wang, X. L., He, X. J., 2016. Anti-inflammatory oleanolic triterpenes from Chinese Acorns. *Molecules* 21.

Ikeda, K., Takahashi, M., Nishida, M., Miyauchi, M., Kizu, H., Kameda, Y., Arisawa, M., Watson, A. A., Nash, R. J., Fleet, G. W. J., Asano, N., 2000. Homonojirimycin analogues and their glucosides from *Lobelia sessilifolia* and *Adenophora* spp. (Campanulaceae). *Carbohyd Res* 323, 73-80.

Ishimaru, K., Nonaka, G., Nishioka, I., 1987. Tannins and related compounds. 57. Gallic acid esters of Protoquercitol, quinic acid and (-)-shikimic acid from *Quercus mongolica* and *Quercus myrsinaefolia*. *Phytochemistry* 26, 1501-1504.

Ito, H., Li, P., Koreishi, M., Nagatomo, A., Nishida, N., Yoshida, T., 2014.

Ellagitannin oligomers and a neolignan from pomegranate arils and their inhibitory effects on the formation of advanced glycation end products. *Food Chem* 152, 323-330.

Jabeen, B., Riaz, N., Saleem, M., Naveed, M. A., Ashraf, M., Alam, U., Rafiq, H. M., Tareen, R. B., Jabbar, A., 2013. Isolation of natural compounds from *Phlomis stewartii* showing α-glucosidase inhibitory activity. *Phytochemistry* 96, 443-448.

Jagetia, G. C., Baliga, M. S., Malagi, K. J., Kamath, M. S., 2002. The evaluation of the radioprotective effect of Triphala (an ayurvedic rejuvenating drug) in the mice exposed to gamma-radiation. *Phytomedicine* 9, 99-108.

Jagtap, A. G., Karkera, S. G., 1999. Potential of the aqueous extract of *Terminalia chebula* as an anticaries agent. *J Ethnopharmacol* 68, 299-306.

Joo, H., Lee, H. J., Shin, E. A., Kim, H., Seo, K. H., Baek, N. I., Kim, B., Kim, S. H., 2016. c-Jun N-terminal kinase-dependent endoplasmic reticulum stress pathway is critically involved in arjunic acid induced apoptosis in non-small cell lung cancer cells. *Phytother Res* 30, 596-603.

Jung, S. W., Shin, M. H., Jung, J. H., Kim, N. D., Im, K. S., 2001. A triterpene glucosyl ester from the roots of *Rubus crataegifolius*. *Arch Pharm Res* 24, 412-415.

Kashiwada, Y., Huang, L., Kilkuskie, R. E., Bodner, A. J., Lee, K. H., 1992a. Anti-aids agents. 5. New hexahydroxydiphenyl derivatives as potent inhibitors of HIV replication in lymphocytes-H9. *Bioorg Med Chem Lett* 2, 235-238.

Kashiwada, Y., Nonaka, G., Nishioka, I., 1984. Tannins and related compounds. 23. Rhubarb. 4. Isolation and structures of new classes of gallotannins. *Chem Pharm Bull* 32, 3461-3470.

Kashiwada, Y., Nonaka, G., Nishioka, I., Chang, J. J., Lee, K. H., 1992b. Antitumor

Agents. 129. Tannins and related compounds as selective cytotoxic agents. *J Nat Prod* 55, 1033-1043.

Kashiwada, Y., Nonaka, G., Nishioka, I., Nishizawa, M., Yamagishi, T., 1988. Studies on Rhubarb (*Rhei Rhizoma*). 14. Isolation and characterization of stilbene glucosides from Chinese Rhubarb. *Chem Pharm Bull* 36, 1545-1549.

Kashtoh, H., Hussain, S., Khan, A., Saad, S. M., Khan, J. A., Khan, K. M., Perveen, S., Choudhary, M. I., 2014. Oxadiazoles and thiadiazoles: novel α-glucosidase inhibitors. *Bioorg Med Chem* 22, 5454-5465.

Kato, A., Minoshima, Y., Yamamoto, J., Adachi, I., Watson, A. A., Nash, R. J., 2008. Protective effects of dietary chamomile tea on diabetic complications. *J Agr Food Chem* 56, 8206-8211.

Khanbabaee, K., van Ree, T., 2001. Tannins: Classification and definition. *Nat Prod Rep* 18, 641-649.

Kim, J. S., Kwon, C. S., Son, K. H., 2000. Inhibition of α-glucosidase and amylase by luteolin, a flavonoid. *Biosci Biotechnol Biochem* 64, 2458-2461.

Klika, K. D., Saleem, A., Sinkkonen, J., Kahkonen, M., Loponen, J., Tahtinen, P., Pihlaja, K., 2004. The structural and conformational analyses and antioxidant activities of chebulinic acid and its thrice-hydrolyzed derivative, 2,4-chebuloyl-β-D-glucopyranoside, isolated from the fruit of *Terminalia chebula*. *Arkivoc*, 83-105.

Konig, M., Scholz, E., Hartmann, R., Lehmann, W., Rimpler, H., 1994. Ellagitannins and complex tannins from *Quercus petraea* Bark. *J Nat Prod* 57, 1411-1415.

Kurokawa, M., Nagasaka, K., Hirabayashi, T., Uyama, S., Sato, H., Kageyama, T., Kadota, S., Ohyama, H., Hozumi, T., Namba, T., Shiraki, K., 1995. Efficacy of

traditional herbal medicines in combination with acyclovir against herpes-simplex virus type-1 infection *in vitro* and *in vivo*. *Antivir Res* 27, 19-37.

Kusano, R., Ogawa, S., Matsuo, Y., Tanaka, T., Yazaki, Y., Kouno, I., 2011. α-Amylase and lipase inhibitory activity and structural characterization of Acacia bark proanthocyanidins. *J Nat Prod* 74, 119-128.

Lee, H. S., Jung, S. H., Yun, B. S., Lee, K. W., 2007. Isolation of chebulic acid from *Terminalia chebula* Retz. and its antioxidant effect in isolated rat hepatocytes. *Arch Toxicol* 81, 211-218.

Li, C. W., Dong, H. J., Cui, C. B., 2015. The synthesis and antitumor activity of twelve galloyl glucosides. *Molecules* 20, 2034-2060.

Li, X. C., Joshi, A. S., ElSohly, H. N., Khan, S. I., Jacob, M. R., Zhang, Z. Z., Khan, I. A., Ferreira, D., Walker, L. A., Broedel, S. E., Raulli, R. E., Cihlar, R. L., 2002. Fatty acid synthase inhibitors from plants: Isolation, structure elucidation, and SAR studies. *J Nat Prod* 65, 1909-1914.

Li Xiang-yu, W. Y.-h., Wang Hong-sheng, Shi Ya-na, Long Chun-lin, 2010. Phenolic derivatives from the leaves of *Dipteronia dyeriana*. Nat Prod Res Dev 22, 5-10.

Luo, H. G., Ma, L., Kong, L. Y., 2008. New triterpenoid saponins with strong α-glucosidase inhibitory activity from the roots of *Gypsophila oldhamiana*. *Bioorgan Med Chem* 16, 2912-2920.

Luyen, B. T. T., Tai, B. H., Thao, N. P., Eun, K. J., Cha, J. Y., Xin, M. J., Lee, Y. M., Kim, Y. H., 2014. Anti-inflammatory components of *Euphorbia humifusa* Willd. *Bioorg Med Chem Lett* 24, 1895-1900.

Machumi, F., Midiwo, J. O., Jacob, M. R., Khan, S. I., Tekwani, B. L., Zhang, J.,

Walker, L. A., Muhammad, I., 2013. Phytochemical, antimicrobial and antiplasmodial investigations of *Terminalia brownii*. *Nat Prod Commun* 8, 761-764.

Madikizela, B., Aderogba, M. A., Finnie, J. F., Van Staden, J., 2014. Isolation and characterization of antimicrobial compounds from *Terminalia phanerophlebia* Engl. & Diels leaf extracts. *J Ethnopharmacol* 156, 228-234.

Mahato, S. B., Kundu, A. P., 1994. ^{13}C NMR spectra of pentacyclic triterpenoids - A compilation and some salient features. *Phytochemistry* 37, 1517-1575.

Marchal, A., Waffo-Teguo, P., Genin, E., Merillon, J. M., Dubourdieu, D., 2011. Identification of new natural sweet compounds in wine using centrifugal partition chromatography gustatometry and fourier transform mass spectrometry. *Anal Chem* 83, 9629-9637.

Marzouk, M. S. A., El-Toumy, S. A. A., Moharram, F. A., Shalaby, N. M. M., Ahmed, A. A. E., 2002. Pharmacologically active ellagitannins from *Terminalia myriocarpa*. *Planta Medica* 68, 523-527.

Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F., Turner, R. C., 1985. Homeostasis model assessment - Insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in Man. *Diabetologia* 28, 412-419.

Mbaze, L. M., Poumale, H. M. P., Wansi, J. D., Lado, J. A., Khan, S. N., Iqbal, M. C., Ngadjui, B. T., Laatsch, H., 2007. α-Glucosidase inhibitory pentacyclic triterpenes from the stem bark of *Fagara tessmannii* (Rutaceae). *Phytochemistry* 68, 591-595.

Mininel, F. J., Leonardo, C. S., Espanha, L. G., Resende, F. A., Varanda, E. A., Leite, C. Q. F., Vilegas, W., dos Santos, L. C., 2014. Characterization and

quantification of compounds in the hydroalcoholic extract of the leaves from *Terminalia catappa* Linn. (Combretaceae) and their mutagenic activity. *Evid-Based Compl Alt*.

Naveed, M. A., Riaz, N., Saleem, M., Jabeen, B., Ashraf, M., Ismail, T., Jabbar, A., 2014. Longipetalosides A-C, new steroidal saponins from *Tribulus longipetalus*. *Steroids* 83, 45-51.

Nawwar, M. A. M., Hussein, S. A. M., Merfort, I., 1994. NMR spectral analysis of polyphenols from *Punica granatum*. *Phytochemistry* 36, 793-798.

Newman, D. J., Cragg, G. M., Snader, K. M., 2000. The influence of natural products upon drug discovery. *Nat Prod Rep* 17, 215-234.

Nonaka, G., Ishimaru K., Mihashi K., Iwase Y., Ageta M., Nishioka I., 1988. Tannins and related compounds. LXIII. Isolation and characterization of mongolicains A and B, novel tannins from *Quercus* and *Castanopsis* species. *Chem. Pharm. Bull.* 36, 857-869.

Okuda, T., Ito, H., 2011. Tannins of constant structure in medicinal and food plants-hydrolyzable tannins and polyphenols related to tannins. *Molecules* 16, 2191-2217.

Okuda, T., Yoshida, T., Ashida, M., Yazaki, K., 1983. Tannins of *Casuarina* and *Stachyurus* species. 1. Structures of pendunculagin, casuarictin, strictinin, casuarinin, casuariin, and stachyurin. *J Chem Soc Perk T* 1, 1765-1772.

Okuda, T., Yoshida, T., Hatano, T., Koga, T., Toh, N., Kuriyama, K., 1982. Circular dichroism of hydrolyzable tannins. I. Ellagitannins and gallotannins. *Tetrahedron Lett* 23, 3937-3940.

Orabi, M. A. A., Taniguchi, S., Yoshimura, M., Yoshida, T., Kishino, K., Sakagami, H., Hatano, T., 2010. Hydrolyzable tannins of Tamaricaceous Plants. III. Hellinoyl-

and macrocyclic type ellagitannins from *Tamarix nilotica*. *J Nat Prod* 73, 870-879.

Pfundstein, B., El Desouky, S. K., Hull, W. E., Haubner, R., Erben, G., Owen, R. W., 2010. Polyphenolic compounds in the fruits of Egyptian medicinal plants (*Terminalia bellerica, Terminalia chebula*, and *Terminalia horrida*): Characterization, quantitation and determination of antioxidant capacities. *Phytochemistry* 71, 1132-1148.

Plumb, R. S., Johnson, K. A., Rainville, P., Smith, B. W., Wilson, I. D., Castro-Perez, J. M., Nicholson, J. K., 2006. UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun Mass Sp* 20, 2234-2234.

Ponnusamy, S., Zinjarde, S., Bhargava, S., Rajamohanan, P. R., RaviKumar, A., 2012. Discovering bisdemethoxycurcumin from *Curcuma longa* rhizome as a potent small molecule inhibitor of human pancreatic α-amylase, a target for type-2 diabetes. *Food Chem* 135, 2638-2642.

Potterat, O., Hamburger, M., 2013. Concepts and technologies for tracking bioactive compounds in natural product extracts: generation of libraries, and hyphenation of analytical processes with bioassays. *Nat Prod Rep* 30, 546-564.

Qi, J.-h. W., Xian; Chen, Gui-lin; Hatano, Tsutomu, 2013. Polyphenolic compounds and antioxidant activity of *Terminalia chebula* retz. var. tomentella Kurt. *Zhongguo Xiandai Zhongyao* 15, 1036-1041.

Raju, D., Ilango, K., Chitra, V., Ashish, K, 2009. Evaluation of anti-ulcer activity of methanolic extract of *Terminalia cheubla* fruits in experimental rats. *J Pharma Sci Res* 1, 101-107.

Ramirez-Espinosa, J. J., Rios, M. Y., Paoli, P., Flores-Morales, V., Camici, G., de la

Rosa-Lugo, V., Hidalgo-Figueroa, S., Navarrete-Vazquez, G., Estrada-Soto, S., 2014. Synthesis of oleanolic acid derivatives: *In vitro*, *in vivo* and *in silico* studies for PTP-1B inhibition. *Eur J Med Chem* 87, 316-327.

Rathinamoorthy, R. a. T., G., 2014. *Terminalia chebula* - Review on pharmacological and biochemical studies. *Pharm Tech* 6, 97-116.

Sabu, M. C., Kuttan, R., 2002. Anti-diabetic activity of medicinal plants and its relationship with their antioxidant property. *J Ethnopharmacol* 81, 155-160.

Saha, P. K., Patra, P. H., Pradhan, R., Dey. S., Mandal, T. K., 2011. Effect of *Terminalia chebula* and *Terminalia bellerica* on wound healing induced dermal wounds in rabbits. *Pharmacology online* 2, 235-241.

Saijo, R., Nonaka, G., Nishioka, I., 1989. Tannins and related compounds. 87. Isolation and characterization of 4 new hydrolyzable tannins from the leaves of *Mallotus repandus*. *Chem Pharm Bull* 37, 2624-2630.

Seger, C., Sturm, S., Stuppner, H., 2013. Mass spectrometry and NMR spectroscopy: modern high-end detectors for high resolution separation techniques - state of the art in natural product HPLC-MS, HPLC-NMR, and CE-MS hyphenations. *Nat Prod Rep* 30, 970-987.

Seo, E. J., Curtis-Long, M. J., Lee, B. W., Kim, H. Y., Ryu, Y. B., Jeong, T. S., Lee, W. S., Park, K. H., 2007. Xanthones from *Cudrania tricuspidata* displaying potent α-glucosidase inhibition. *Bioorg Med Chem Lett* 17, 6421-6424.

Shai, L. J., Magano, S. R., Lebelo, S. L., Mogale, A. M., 2011. Inhibitory effects of five medicinal plants on rat α-glucosidase: Comparison with their effects on yeast α-glucosidase. *J Med Plants Res* 5, 2863-2867.

Sivasothy, Y., Loo, K. Y., Leong, K. H., Litaudon, M., Awang, K., 2016. A potent

α-glucosidase inhibitor from *Myristica cinnamomea* King. *Phytochemistry* 122, 265-269.

Tabopda, T. K., Ngoupayo, J., Liu, J., Mitaine-Offer, A. C., Tanoli, S. A. K., Khan, S. N., Ali, M. S., Ngadjui, B. T., Tsamo, E., Lacaille-Dubois, M. A., Luu, B., 2008. Bioactive aristolactams from *Piper umbellatum*. *Phytochemistry* 69, 1726-1731.

Tanaka, T., Fukumori, M., Ochi, T., Kouno, I., 2003. Paeonianins A-E, new dimeric and monomeric ellagitannins from the fruits of *Paeonia lactiflora*. *J Nat Prod* 66, 759-763.

Tanaka, T., Nakashima, T., Ueda, T., Tomii, K., Kouno, I., 2007. Facile discrimination of aldose enantiomers by reversed-phase HPLC. *Chem Pharm Bull* 55, 899-901.

Tanaka, T., Nonaka, G., Nishioka, I., 1983. Tannins and related compounds. 14. 7-*O*-galloyl-(+)-catechin and 3-*O*-galloylprocyanidin-B3 from *Sanguisorba officinalis*. *Phytochemistry* 22, 2575-2578.

Tanaka, T., Nonaka, G., Nishioka, I., 1986a. Tannins and related compounds. 41. Isolation and characterization of novel ellagitannins, punicacorteins A, B, C and D, and punigluconin from the Bark of *Punica granatum* L. *Chem Pharm Bull* 34, 656-663.

Tanaka, T., Nonaka, G. I., Nishioka, I., 1986b. Tannins and related compounds. 42. Isolation and characterization of 4 new hydrolyzable tannins, terflavin A and terflavin B, tergallagin and tercatain from the leaves of *Terminalia catappa* L. *Chem Pharm Bull* 34, 1039-1049.

Tasduq, S. A., Singh, K., Satti, N. K., Gupta, D. K., Suri, K. A., Johri, R. K., 2006. *Terminalia chebula* (fruit) prevents liver toxicity caused by sub-chronic

administration of rifampicin, isoniazid and pyrazinamide in combination. *Hum Exp Toxicol* 25, 111-118.

Uvarani, C., Jaivel, N., Sankaran, M., Chandraprakash, K., Ata, A., Mohan, P. S., 2014. Axially chiral biscarbazoles and biological evaluation of the constituents from *Murraya koenigii*. *Fitoterapia* 94, 10-20.

Verzele, M., Delahaye, P., Vandijck, J., 1983. Digallic Acid. *B Soc Chim Belg* 92, 181-186.

Walia, H. a. A., S., 2013. *Terminalia chebula* - A pharmacognistic account. *J Med Plants Res* 7, 1351-1361.

Wilkins, C., 1988. Galloyl glucose derivatives from *Heuchera cylindrica*. *Phytochemistry* 27, 2317-2318.

Yang, B. R., Kortesniemi, M., Liu, P. Z., Karonen, M., Salminen, J. P., 2012. Analysis of hydrolyzable tannins and other phenolic compounds in emblic leaf flower (*Phyllanthus emblica* L.) fruits by high performance liquid chromatography-electrospray ionization mass spectrometry. *J Agr Food Chem* 60, 8672-8683.

Yang, S. W., Zhou, B. N., Wisse, J. H., Evans, R., van der Werff, H., Miller, J. S., Kingston, D. G. I., 1998. Three new ellagic acid derivatives from the bark of *Eschweilera coriacea* from the Suriname rainforest. *J Nat Prod* 61, 901-906.

Yoshida, T., Fujii, R., Okuda, T., 1980. Revised structures of chebulinic acid and chebulagic acid. *Chem Pharm Bull* 28, 3713-3715.

Yoshida, T., Itoh, H., Matsunaga, S., Tanaka, R., Okuda, T., 1992. Tannins and related polyphenols of Euphorbiaceous plants. 9. Hydrolyzable tannins with $^{1}C_{4}$ glucose core from *Phyllanthus flexuosus* Muell. *Chem Pharm Bull* 40, 53-60.

Yoshida, T., Okuda, T., Koga, T., Toh, N., 1982. Absolute configurations of

chebulic, chebulinic and chebulagic acid. *Chem Pharm Bull* 30, 2655-2658.

Yoshida, T., Okuda, T., Memon, M. U., Shingu, T., 1985. Tannins of rosaceous medicinal plants. 2. Gemin A, Gemin B, and Gemin C, new dimeric ellagitannins from *Geum japonicum*. *J Chem Soc Perk T* 1, 315-321.

Yukawa, T. A., Kurokawa, M., Sato, H., Yoshida, Y., Kageyama, S., Hasegawa, T., Namba, T., Imakita, M., Hozumi, T., Shiraki, K., 1996. Prophylactic treatment of cytomegalovirus infection with traditional herbs. *Antivir Res* 32, 63-70.

Zhang, H. M., Wang, C. F., Shen, S. M., Wang, G. L., Liu, P., Liu, Z. M., Wang, Y. Y., Du, S. S., Liu, Z. L., Deng, Z. W., 2012. Antioxidant phenolic compounds from Pu-erh tea. *Molecules* 17, 14037-14045.

Zhang, Y. J., Abe, T., Tanaka, T., Yang, C. R., Kouno, I., 2001. Phyllanemblinins A-F, new ellagitannins from *Phyllanthus emblica*. *J Nat Prod* 64, 1527-1532.

Zhang, Y. J., DeWitt, D. L., Murugesan, S., Nair, M. G., 2004. Novel lipid-peroxidation- and cyclooxygenase-inhibitory tannins from *Picrorhiza kurroa* seeds. *Chem Biodivers* 1, 426-441.

Zhou, X. H., Kasai, R., Ohtani, K., Tanaka, O., Nie, R. L., Yang, C. R., Zhou, J., Yamasaki, K., 1992. Oleanane and ursane glucosides from *Rubus* species. *Phytochemistry* 31, 3642-3644.

국문초록

가자나무 (*Terminalia chebula* Retz.)는 사군자과 (Combretaceae) 식물로써 주로 인도, 동남아시아, 중국남부 등에 분포하고 있다. 오래전부터 말린 가자나무의 열매는 약재로 사용되어 왔으며, 완하, 강장, 건위, 지혈, 항 말라리아 등 다양한 목적으로 사용되어 왔다. 이러한 특성으로 인해 가자나무 열매의 화학적 성분은 오랜기간 연구되어 왔으며, 지금까지 hydrolysable tannin, phenolic acid, triterpene, flavonoid, lignan과 그 배당체를 포함한 약 150여종의 다양한 성분들이 분리, 보고되어 있다.

천연물은 오랜기간 신약의 보고로 활용되어 왔으며 그 중요성은 현재에도 널리 인정받고 있다. 하지만 천연물의 연구 방법 중 가장 널리 사용되는 bioactivity-guided isolation 기법은 비용과 시간, 노동력이 많이 소모되며 기존에 분리 보고 된 화합물을 다시 분리할 수도 있는 단점을 가지고 있다. 이러한 단점을 보완하기 위해 많은 연구들이 진행되었는데, 그 중 HPLC를 이용하여 활성 추출물 및 분획물을 96 well plate에 직접 fractionation하고 이에 대한 활성 평가를 통해 간단하게 활성 물질을 확인할 수 있는 HPLC-based activity profiling 기법이 주목을 받고 있다. 또한 이 기법은 HRMS, NMR 등의 고해상도 기기를 이용한 dereplication법을 쉽게 적용할 수 있어, 기존에 분리된 물질과 신규화합물의 여부를 확인할 수 있다. 본 연구에서는 HPLC-based activity profiling과 dereplication법을 이용하여 Baker's yeast α-glucosidase에 대한 억제 활성이 있는 물질을 규명해 보았다.

가자나무 열매의 메탄올 추출물 및 n-hexane, $CHCl_3$, n-BuOH, 물

분획물을 대상으로 각각 α-glucosidase 억제 활성을 평가하였으며, 그 결과 n-BuOH층의 활성이 가장 뛰어나다는 것을 확인하였다. 또한 n-BuOH 층을 HP-20 diaion resin을 이용하여 추가로 네 가지 분획 (B1-B4)으로 나누었고 이에 대해 Baker's yeast α-glucosidase 억제 활성을 평가한 결과 B2, B3 분획이 가장 억제 활성이 뛰어나다는 것을 확인할 수 있었다. 그러므로 B2, B3분획을 대상으로 HPLC-based activity profiling 기법과 derelication법을 이용하여 Baker's yeast α-glucosidase 억제 물질의 규명을 시도하였다. 그 결과 총 13종의 화합물을 분리하였으며, 이 중 6′-*O*-methyl neochebulagate (**3**), 4-*O*-(2′′,4′′-di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**11**), 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**), 1,2,3-tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (**13**)의 4종의 신규화합물을 분리할 수 있었다. 또한 13종의 화합물에 대하여 Baker's yeast α-glucosidase 억제 활성을 평가한 결과, 화합물 **1-3**을 제외한 화합물 **4-13** 에서 모두 억제활성이 확인 되었고, 그 중 신규화합물인 화합물 **11** (IC_{50} 6.4 μM), **12** (IC_{50} 2.9 μM)이 가장 뛰어난 억제 활성을 보였다. 이를 통해 HPLC-based activity profiling기법과 derepliation 법의 활용이, 신규 활성 물질 탐색에 효과적임을 확인할 수 있었다.

또한 추가적인 가자열매의 성분연구를 위해 추가적으로 분리를 진행한 결과, 기존에 분리된 13종의 물질을 포함하여 hydrolysable tannin 계열 물질 49종과 polyhydroxytriterpene 계열 물질 17종, 총 66종의 물질을 분리하였고, 이 중 hydrolysable tannin 계열에서 9종, triterpene 계열에서 3종의 신규 화합물을 분리하였다. 이는 1D and 2D NMR (^{1}H, ^{13}C, ^{1}H-^{1}H COSY, edited-HSQC, HMBC, ROESY)과

가수분해, CD spectrum 측정을 통해 각각 6′-*O*-methyl neochebulagate (**3**), 4-*O*-(2′′,4′′-di-*O*-galloyl-*α*-L-rhamnosyl)ellagic acid (**11**), 1,2,3,6-tetra-*O*-galloyl-4-*O*-cinnamoyl-*β*-D-glucose (**12**), 1,2,3-tri-*O*-galloyl-6-*O*-cinnamoyl-*β*-D-glucose (**13**) methyl chebulagate (**18**), 1′-*O*-methyl neochebulanin (**21**), dimethyl noechebulinate (**22**), dimethyl neochebulagate (**23**), dimethyl 4′-*epi*-neochebulagate (**24**) (hydrolysable tannin 계열), 23-*O*-neochebuloylarjungenin 28-*O*-*β*-D-glucopyranosyl ester (**51**), 23-*O*-4′-*epi*-neochebuloylarjungenin (**52**), 23-*O*-galloylpinpaenoic acid 28-*O*-*β*-D-glucopyranosyl ester (**66**) (triterpene 계열)로 확인하였다. 이 중 화합물 **24**, **52**는 neochebulic acid (**37**)에서 나온 4′-*epi*-neochebuloyl기를 분자 내 포함하는 구조로 자연계에서 매우 드물게 발견되는 구조이다.

분리된 화합물을 대상으로 Baker's yeast α-glucosidase에 대한 억제 활성을 평가한 결과 hydrolysable tannin 계열에서 21종 (**4**, **5**, **6**, **8**, **9**, **10**, **11**, **12**, **13**, **16**, **17**, **18**, **22**, **25**, **43**, **44**, **45**, **46**, **47**, **48**, **49**)의 물질이, triterpene 계열에서 2종 (**60**, **61**)의 물질이 억제 활성이 있는 것으로 확인되었으며, 앞선 HPLC-based activity profiling의 결과와 마찬가지로 화합물 **11**, **12**가 가장 뛰어난 억제 활성을 보였다. 그리고 11, 12번 화합물에 대하여 억제 모드를 결정하기 위해 Lineweaver-Burk double-reciprocal plot을 이용하여 분석한 결과 두 화합물 모두 Mixed inhibition mode를 가지는 것으로 확인되었다.

또한 분리된 화합물을 대상으로 추가적으로 rat intestinal α-

glucosidase, porcine pancreatic α-amylase, protein-tyrosine phosphatase 1B (PTP-1B)에 대하여 억제 활성을 확인하였다. 이 중, rat intestinal α-glucosidase, porcine pancreatic α-amylase에 대해서는 양성대조군인 acarbose에 비하여 1,2,3,4,6-penta-*O*-galloyl-*β*-D-glucose (**8**) 만이 억제 활성을 나타내었다. PTP-1B에 대해서는 화합물 **5**, **6**, **8**, **11**, **12**, **13**, **48**, **49**, **53**, **56**, **58**, **60**이 IC_{50} 20 μM 이하의 값을 가지는 것으로 확인되었다.

주요어: 가자나무, 사군자과, hydrolysable tannin, polyhydroxytriterpene, HPLC-based activity profiling, α-glucosidase 억제활성

학번: 2009-21699